# NAVAL POSTGRADUATE SCHOOL
## Monterey, California

AD-A285 512

|||||||||||||||||||||||||| 2

# THESIS

OCT 1 1994

**A COST SIMULATION TOOL FOR ESTIMATING THE COST OF OPERATING GOVERNMENT OWNED AND OPERATED SHIPS**
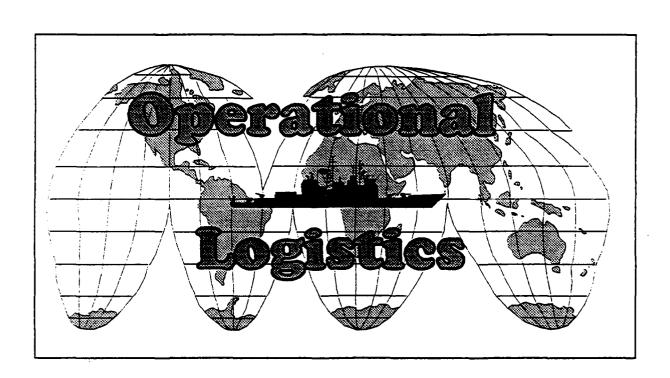
Terry Redman

September, 1994

Thesis Co-Advisors:                         Dan C. Boger
                                            William G. Kemple

Approved for public release; distribution is unlimited

# REPORT DOCUMENTATION PAGE

Form Approved OMB Np. 0704

| 1. AGENCY USE ONLY *(Leave blank)* | 2. REPORT DATE September 1994 | 3. REPORT TYPE AND DATES COVERED Master's Thesis |
|---|---|---|

| 4. TITLE AND SUBTITLE A COST SIMULATION TOOL FOR ESTIMATING THE COST OF OPERATING GOVERNMENT OWNED AND OPERATED SHIPS | 5. FUNDING NUMBERS |
|---|---|
| 6. AUTHOR(S) Redman, Terry L. | |

| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey CA 93943-5000 | 8. PERFORMING ORGANIZATION REPORT NUMBER |
|---|---|

| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) | 10. SPONSORING/MONITORING AGENCY REPORT NUMBER |
|---|---|

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

| 12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited. | 12b. DISTRIBUTION CODE |
|---|---|

13. ABSTRACT *(maximum 200 words)*

The cost of operating ships is difficult to predict. A historic ship's operating cost detabase is maintained by the Military Sealift Command(MSC); but, it is very difficult to extract or manipulate the data to support prediction or regression analysis. An alternative was sought that would reduce the effort for the user when attempting to make predictions from the data. If the data for each cost category (salary, training, fuel, port and miscellaneous, subsistence, ship's equipage, and voyage repairs) could be well approximated using probability distributions, then the costs of an operational scenario, with estimates of the uncertainties, could be obtained through use of a Monte Carlo simulation.

The MSC data was divided into subsets, one for model fitting and one for validation. Once probability distributions had been fit to the data, a Monte Carlo simulation tool was developed using the Crystal Ball simulation add in to Microsoft Excel. The data analysis and cost model were then validated using empirical data.

Based on the results, the Cost Simulation model provides a useful tool for predicting operating costs and supports sensitivity analysis of various ship's operating cost scenarios.

| 14. SUBJECT TERMS    Monte Carlo Simulation, Data Analysis | 15. NUMBER OF PAGES   150 |
|---|---|
| | 16. PRICE CODE |

| 17. SECURITY CLASSIFICATION OF REPORT Unclassified | 18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified | 19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified | 20. LIMITATION OF ABSTRACT UL |
|---|---|---|---|

A Cost Simulation Tool for Estimating the Cost of Operating Government
Owned and Operated Ships

by

Terry Lee Redman
Lieutenant Commander, United States Navy
B.S., United States Merchant Marine Academy, 1980

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 1994

Author: _____
Terry Lee Redman

Approved by: _____
Dan Boger, Thesis Co-Advisor

_____
William Kemple, Thesis Co-Advisor

_____
Shu S. Liao, Second Reader

_____
Peter Purdue, Chairman
Department of Operations Research

ii

# ABSTRACT

The cost of operating ships is difficult to predict. A historic ship's operating cost database is maintained by the Military Sealift Command (MSC); but, it is very difficult to extract or manipulate the data to support prediction or regression analysis. An alternative was sought that would reduce the effort for the user when attempting to make predictions from the data. If the data for each cost category (salary, training, fuel, port and miscellaneous, subsistence, ship's equipage, and voyage repairs) could be well approximated using probability distributions, then the costs of an operational scenario, with estimates of the uncertainties, could be obtained through use of a Monte Carlo simulation.

The MSC data was divided into two subsets, one for model fitting and one for validation. Once probability distributions had been fit to the data, a Monte Carlo simulation tool was developed using the Crystal Ball® simulation add in to Microsoft Excel®. The data analysis and cost model were then validated using the empirical data.

Based on the results, the Cost Simulation model provides a useful tool for predicting operating costs and supports sensitivity analysis of various ship's operating cost scenarios.

## THESIS DISCLAIMER

The reader is cautioned that computer programs developed in this research may not have been exercised for all cases of interest. While every effort has been made, within the time available, to ensure that the programs are free of computational and logic errors, they cannot be considered validated. Any application of these programs without additional verification is at the risk of the user.

## ACKNOWLEDGMENT

| Accesion For | |
|---|---|
| NTIS CRA&I | ☑ |
| DTIC TAB | ☑ |
| Unannounced | ☐ |
| Justification | |
| By | |
| Distribution / | |
| Availability Codes | |
| Dist | Avail and / or Special |
| A-1 | |

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# LIST OF FIGURES

# LIST OF FIGURES

## EXECUTIVE SUMMARY

The cost of operating ships is difficult to predict. A historic ship's operating cost database is maintained by the Military Sealift Command (MSC); but, it is very difficult to extract or manipulate the data to support prediction or regression analysis. An alternative was sought that would reduce the effort for the user when attempting to make predictions from the data. If the data for each cost category (salary, training, fuel, port and miscellaneous, subsistence, ship's equipage, and voyage repairs) could be well approximated using probability distributions, then the costs of an operational scenario, with estimates of the uncertainties, could be obtained through use of a Monte Carlo simulation.

The MSC data was divided into two subsets, one for model fitting and one for validation. The model fitting subset was analyzed using graphical data analysis. The object of this effort was to fit probability distributions to the data. Good probability distribution fits were found for each of the cost categories examined in the subset.

Once probability distributions had been fit to the data, a Monte Carlo simulation tool was developed using the Crystal Ball® simulation add-in to Microsoft Excel®. The simulation tool was designed with a user interface to reduce the technical knowledge needed by the user to operate the application.

The data analysis and cost model were then validated using the empirical data. The simulation tool was run and the results compared to the actual per diem rates in use at MSC. This analysis showed that the simulation model produced results close to the actual per diem rate. The results pointed to either possible inaccuracies in the overhead rates used at MSC, or problems with the simulation model.

The simulation model was run without indirect (overhead) costs included and compared to the historical direct costs in the entire database. These runs showed the simulation model to accurately predict direct costs.

On the basis of the results, the Cost Simulation model provides a useful tool for predicting direct operating costs and supports sensitivity analysis of various ship's operating cost scenarios. Further study is required in the area of indirect (overhead) costs to permit use of the simulation model for prediction of total costs of operation. This would enable MSC to use the simulation tool for setting per diem rates with a higher degree of accuracy.

# I. BACKGROUND AND PROBLEM STATEMENT

Military Sealift Command, United States Pacific Fleet (MSCPAC) is a civilian manned and operated shipping line owned by the United States Government. The mission of MSCPAC is fourfold:

1. Operate and maintain Naval Fleet Auxiliary Force (NFAF) ships to provide direct support to U.S. Navy combatant ships.

2. Operate and maintain Prepositioned Sealift ships as required by the National Command Authority (NCA).

3. Operate and maintain Special Missions Force ships as required for specialized military purposes such as oceanographic and hydrographic surveys, undersea surveillance, and missile telemetry collection and range instrumentation.

4. Charter and contract ocean cargo services as necessary to support the military commitments.

The first three of the missions are performed primarily by civilian crewed ships that are owned and operated by the government. To accomplish this, MSCPAC operates a fleet of thirty-four ships at an annual cost of $500 million.

In recent years, the funding scheme was changed from the Industrial Fund to the Defense Business Operating Fund (DBOF). Under DBOF funding, the sponsor for each ship is billed for the services provided by the

ship. The charges are billed on a cost per day or "per diem" basis. The sponsor pays these bills using his Operating and Maintenance, Navy Funds (OMN). MSCPAC must meet all of its operating costs using funds obtained in this manner.

In the post cold-war atmosphere of dwindling defense budgets, the sponsors are in some cases demanding a lower per-diem charge. In the case of Special Mission ships, the sponsor of the cable ships, Navy Space Surveillance and Warfare Command (SPAWAR), has threatened to contract with a civilian contractor who is bidding a lower price.

Any overcharges reduce the precious OMN funds that are needed to conduct all the sponsor's missions. Eventually, when the ship sponsor's budget is cut further, he will be forced to fulfill his requirements with the lowest bid, which frequently is the civilian contractor. This increased need for cost competitiveness requires the ability to forecast budget needs to a greater degree of accuracy than is presently common in government practice. It is clearly in the best interest of MSCPAC to preserve their viability as an economical alternative to the civilian contractors by obtaining the ability to make tighter budget forecasts. The costs of operation be analyzed and cut to the minimum to permit MSCPAC to quote the lowest per-diem rate possible.

There are three major categories of cost in the MSCPAC operation, the cost of overhead for facilities and shore based employees, the cost of

chartering and contracting cargo services, and the cost of operating the government owned and operated ships. In the opinion of the MSCPAC comptroller and staff, the major difficulties are in the third category.

The first category should be fairly predictable and should not contribute greatly to the total cost of MSCPAC.

The second category comprises a different form of control by MSCPAC. During the negotiating of a charter or contract for services, the best price is sought. When finally billed for the charters and contracts, MSCPAC pays the predetermined price. MSCPAC is later reimbursed by the sponsor for the services provided by MSCPAC through the charter or contract. This area normally does not present a problem to the accounting department.

The final category, the cost of operating the government owned and operated ships is the area of most concern to the comptroller.

But with existing data and tools, questions from management for cost analysis information become painful forays into voluminous heaps of data that are difficult to interpret and analyze. Often, the answers require several man days effort to collate and assemble the data into a manageable form. The results are not timely enough to meet the demands for information on which to base decisions.

Presently, all payment vouchers written by MSC are paid by MSC Headquarters in Washington, DC. The transactions are recorded on a

mainframe based accounting system called the Financial Management Information System (FMIS). The area headquarters such as MSCPAC can access data from FMIS through a personal computer landline linkup system called FMIS Gateway.

The PC/mainframe interface is cumbersome. The information is accessible on-line, but it is very difficult to analyze one screen at a time. All reports are routinely downloaded from FMIS Gateway in the form of printouts. When the need for analysis arises, the accounting department must gather vast amounts of paper records. Frequently, by the time an answer is produced, the question is no longer of interest!

Management needs a more responsive tool for budgeting and conducting "what if" analysis. This tool should aid in budget forecasting by allowing management to predict the effects of anticipated fiscal changes with a reasonable amount of accuracy. The ability to conduct sensitivity analysis would permit the tightening of costs in areas that would most affect the total cost of operation.

Another advantage of a tool of this type would be the ability to analyze year to date expenditures. If certain fiscal policies can be predicted to outspend the budget before the fact, those practices can be changed to allow the activity to remain within budget. At present, this ability is almost non-existent other than the traditional "stubby pencil" methods.

The purpose of this thesis is to develop and demonstrate a cost simulation tool through which this process can be streamlined so answers can be found within a matter of minutes rather than days. The tool is based on the premise that each of the categories of operating cost for the government owned and operated ships behaves in a way that can be modeled using known probability distributions. Once these probability distributions are known, cost studies for different periods of times and conditions can be made very quickly using a personal computer and Monte Carlo simulation.

The probability and simulation models are combined in a popular spreadsheet program with an add-in simulation solver. This powerful combination brings the ability to make decisions based on solid forecasting directly to the manager's desktop. No longer will the manager be totally at the mercy of the accounting department when he needs rapid answers to his questions.

The remainder of this thesis is arranged as follows:

1. Chapter II explains the methodology used in the project.

2. Chapter III gives a description of the components of each cost category.

3. Chapter IV presents the data analysis. The chapter is rather lengthy and is summarized in the first section. A more detailed reading of the data analysis is the subject of the remainder of the chapter.

4. Chapter V explains the cost analysis model.

5. Chapter VI briefly explains the Monte Carlo simulation method as applied in this cost simulation model.

6. Chapter VII presents a validation of the data analysis.

7. Chapter VIII gives the conclusions.

8. The appendix gives the visual basic code used in building the enhanced user interface for the spreadsheet model.

## II. METHODOLOGY

Exploratory data analysis and Monte Carlo simulation are combined to build a cost simulation tool to predict the direct costs of operating the MSC owned and operated ships. Monte Carlo simulation is a scheme in which random numbers are used for solving certain stochastic or deterministic problems where the passage of time plays no substantive role. In this case, random numbers from a uniform distribution on the unit interval, $U(0,1)$, will be processed through probability distribution transformations to model the operating cost category data according to relationships found by performing data analysis on the historical operating cost data.

The Monte Carlo simulation model is implemented as an add-in to the spreadsheet commonly used at MSCPAC, Microsoft Excel®. The Monte Carlo simulation is implemented in Crystal Ball® from Decisioneering Inc., Denver, Colorado, which runs as an add-in to Microsoft Excel. This system runs on any PC that is capable of running Microsoft Excel. In most cases at MSCPAC, the software will be run on a 386 PC.

This approach requires an extensive data analysis of historical cost data. The data is in the form of monthly expense records and is analyzed to obtain probability distributions to model each category for each ship class.

The central assumptions behind the cost simulation tool are that the ships operating costs are random in nature and can be modeled by known probability distributions.

A spreadsheet database is first created in Microsoft Excel to enable the transfer of data to the data analysis program, *A Graphical Statistical System* (AGSS). *A Graphical Statistical System* (AGSS) is used at the Naval Postgraduate School under a test site agreement with IBM Research. We are indebted to Dr. Peter Welch for making this possible.

Due to software incompatibility between the FMIS Gateway system and the PC spreadsheets, the data was hand keyed into the database. Future enhancement of the cost simulation tool could be achieved by establishing direct links between FMIS Gateway and the spreadsheet database.

The actual data analysis is accomplished in AGSS. Each ship class cost category becomes a variable in AGSS. These variables are first analyzed using descriptive plots such as histograms and kernel density estimates to determine which theoretical distributions might be appropriate. The variables are next plotted and analyzed using the probability distribution fitting capability of AGSS. Each variable is plotted in various manners and goodness-of-fit statistics are calculated to determine the best probability distribution fit. Figure 1 depicts the three plot view used for probability distribution fitting.

## T-ATF 166 Class Voyage Repair Cost
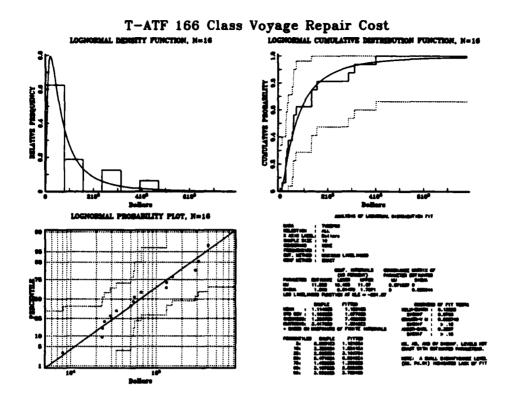


**Figure 1**
**Sample Probability Distribution Plot**

In the upper left hand plot, the data is plotted as a histogram of relative frequency of occurrence versus cost. This histogram is superimposed upon a plot of the selected probability density which has been generated by AGSS. AGSS computes the maximum likelihood estimates of the distribution parameters, together with asymptotic confidence regions.

The upper right plot shows the data set plotted as an empirical cumulative frequency distribution superimposed on the cumulative distribution function for the fitted probability distribution. In this view,

9

AGSS offers the ability to plot Kolmogorov-Smirnov bounds as a further aid in determining a good fit between the data and the theoretical distribution.

The lower left hand plot in Figure 1 shows a probability plot. In this plot, the theoretical distribution is represented as a straight line diagonally across the plot. The data is plotted and superimposed such that if it was a perfect fit to the theoretical distribution, the plot would fall exactly on the line.

The lower right hand section of Figure 1 contains tabular data that helps determine goodness of fit. Four different goodness of fit tests are provided for uncensored data. These are the Chi-Square, Kolmogorov-Smirnov, Cramer-von Mises, and Anderson-Darling tests. Due to the size of the figures, the tabular data for probability distribution plots used in this analysis is found in Appendix A.

By analyzing the plots and comparing them with those of other distributional fits, it is possible to determine the best distributional fit for the data. These distributional fits are used as distributional assumptions in the simulation model.

The cost simulation model is written in the form of a Microsoft Excel workbook. Each worksheet page in the workbook is used for a different ship class or model variation. The ship classes examined in this project will be the T-AO 187 class fleet oiler and T-ATF 166 class oceangoing tug.

10

Microsoft Excel is used primarily for the convenience of the intended users. MSCPAC personnel are familiar with Microsoft Excel so use of Excel for this model will reduce the training requirements for the users.

Microsoft Excel supports the use of the Crystal Ball Monte Carlo simulation add-in. This product is available in versions for Microsoft Excel and Lotus 1,2,3. Crystal Ball is a user-friendly, graphically-oriented forecasting and risk analysis program. Through Monte Carlo simulation, Crystal Ball forecasts the entire range of results possible for a given situation. Crystal Ball uses probability distributions to describe the uncertainty in the assumption cells of the spreadsheet model. There are sixteen different probability distributions available in Crystal Ball for describing the relationships in the model.

Microsoft Excel allows the user to operate in the Microsoft Windows environment. The Windows graphical user interface (GUI) allows users to quickly grasp complex concepts in a more intuitive manner. All planned cost simulation tool users are already skilled in using Microsoft Windows based programs.

Microsoft Excel also supports a rich macro language (Visual Basic) to allow automating the application. This macro language is used to develop a simplified interface for the users. The primary goals for the cost simulation tool are accurate results and ease of use.

Pairwise correlation between each cost category is examined using the Spearman Rank correlation coefficient. Where there is any noted pairwise correlation, it is compensated for in the assumption parameters of Crystal Ball. Multiple correlation is also examined.

The final facet in establishing the accuracy and creditability of the cost simulation tool is to thoroughly validate the model and the distributional assumptions. The first method of validation is to divide the data and analyze each subset independently to determine whether distributional assumptions are valid for all subsets of the data set. In this case, the first half of the data set (chronologically divided) is used to obtain initial assumptions and the second chronological half of the data is used to validate the assumptions.

Finally, test run simulations are conducted to determine the accuracy of the cost simulation that corresponds to the actual conditions for randomly selected ships and months. The results of the simulations are compared to the actual month end account figures.

# III. MILITARY SEALIFT COMMAND COSTS

MSC accounts for sh.. operating costs using direct costs and indirect costs. The direct costs are made up of cost categories that are economically feasible to trace to a particular ship (or cost center). Indirect costs are overhead costs that are not economically feasible or practical to trace to an MSC ship. For example, the cost of xerox copy paper for use in the headquarters offices are paid for by DBOF revenues from the ships but its consumption cannot be directly traced to a particular ship.

Indirect costs are difficult to accurately predict. Indirect costs from previous years are examined and a budget estimate is made based on the amounts spent in previous years. An indirect cost budget for the coming fiscal year is then allocated as a percentage of estimated direct costs for the year. The total cost charged to a sponsor is simply:

*Total Cost = Direct Costs + Allocated Indirect Costs*

This analysis focused only on the direct costs in an effort to show a relationship between the historical database and known probability distributions. This permits the use of Monte Carlo simulation for predicting direct costs.

The overhead costs (designated as administrative indirect costs for this analysis) were not precisely known by MSC and were not directly related to the actual operation of the ships. The Cost Simulation Tool cannot simulate these costs since one of its assumptions is that costs are directly related to the operation of the ships. The overhead rate used in the model was given by MSC. No attempt was made by this analysis to simulate these costs.

Overhead costs are allocated as a percentage of direct cost throughout the Department of Defense[1]. This method is also used at MSC. An overhead rate of 19% of the direct cost was given by the MSCPAC Director of Operations as the current rate being used by MSC. The calculation of the rate is as follows:

$$Overhead\ Rate = (Budgeted\ Indirect\ Costs \div Estimated\ Direct\ Costs) \times 100$$

## A. INDIRECT COSTS.

Three categories of indirect cost are allocated to two overhead cost pools. The first pool of overhead costs used will be known as the administrative pool and includes the following:

1. Headquarters and other administrative support costs.
2. Physical plant and building costs.

---

[1]Young, Douglas, "Complexities, Impact of Overhead", U.S. Army Comptroller Office, Resource Management, 1st Quarter, 1994

This pool makes up about 65% of the model's budgeted indirect cost or about 19% of the actual direct costs. This pool consists all of what would be referred to as overhead costs by MSC.

The second pool of indirect costs used in the model consists of planned maintenance and docking costs. This pool is used in the Cost Simulation Tool model but in actual practice these costs are budgeted direct costs because they are traceable to a particular ship and are budgeted in advance. The model assumption that a ship is 100% available to the sponsor conflicts with the actual case (i.e., that the ships are not available to the sponsor while undergoing planned maintenance and/or drydocking.) The probability distribution modelling feature of the Cost Simulation Tool is only accurate for cases in which the ship is 100% available to the sponsor, therefore the maintenance and docking costs will be treated as if they are an indirect cost. The planned maintenance costs are budgeted in advance and included in the estimate of total direct cost for the coming year. The estimated budgeted overhead rate (administrative indirect cost pool in the model) and the estimate of total direct cost (which includes the planned maintenance and drydocking costs) are used in determining the per diem rate that is charged to the sponsors for use of the ships. Since MSC lumps the costs together in their estimate for per diem rate determination, it is reasonable for this

analysis to model the maintenance and drydocking costs as budgeted indirect costs which will be summed to the direct costs in the model.

No effort was made to obtain probability distributions to fit the planned maintenance and drydocking costs. The historical maintenance costs were summed and allocated as an average percent of total direct cost for each class of ships for the entire period of time for which the data was obtained. For each model run, this allocated percentage is added to the percentage used for the other cost pool. The maintenance and docking costs make up about 35% of the model's budgeted indirect costs and 10.4% of the total direct costs. Since these costs are planned for and known in advance, they are not treated as a random variable. This slight departure from actual practice will permit us to better use the model to predict the overall operating costs.

The budgeted indirect costs will be allocated using percentage of total direct cost as the basis. In actual MSC accounting practice, the overhead costs are also allocated as a percentage of total direct cost.

## B. SALARY COSTS.

This category includes all payroll costs incurred in paying seagoing employees. The following is a detailed list of salary costs as listed on actual FMIS printouts:

1. Salary data communication.
2. Civilian mariner base pay.
3. Civilian mariner overtime pay.
4. Civilian mariner premium pay.
5. Civilian mariner hazardous duty pay.
6. Civilian mariner beneficial suggestion awards.
7. Civilian mariner incentive pay.
8. Civilian mariner awaiting assignment pay.
9. Civilian mariner indoctrination pay.
10. Military pay.
11. Civilian mariner annual leave earned.
12. Civilian mariner sick pay earned.
13. Civilian mariner shore leave earned.
14. Civilian mariner health insurance.
15. Civilian mariner life insurance.
16. Civilian mariner retirement fund.
17. Civilian mariner FICA.

## C. TRAINING COSTS.

This category includes damage control and safety training costs

obtained in route to the ship by officers and crew. The following is a detailed

list of training costs:

1. Officers damage control instructor school.
2. Civilian mariner firefighting training enroute.
3. Civilian mariner small arms training.
4. Civilian mariner safety training.
5. Maritime academy cadet training.
6. Miscellaneous training.
7. Closed circuit television system training.

## D. FUEL COSTS.

This category includes the cost of diesel fuel and petroleum lubricants

and greases.

## E. SUBSISTENCE COSTS.

This category includes the cost of subsistence for civilian officers and crew. Military officer and crew subsistence is subsidized by the military pay account and is not a cost to MSC.

## F. PORT AND MISCELLANEOUS COSTS.

This category includes the cost a number of items. The following is a detailed listing of port and miscellaneous costs:

1. Transportation of items to and from ship.
2. Consumables.
3. Spare parts.
4. ADP supplies.
5. Software.
6. Medical supplies.
7. ADP equipment.
8. Docking and other fees.
9. Piloting and towage.
10. Panama canal tolls.
11. Utilities.
12. Security guards.
13. Civilian mariner repatriation travel.
14. Civilian mariner other travel.
15. Laundry.
16. Movies/tapes.
17. INMARSAT.
18. Medical expenses.
19. Other miscellaneous expenses.

## G. SHIP'S EQUIPAGE

This category includes the cost of equipage items such as binoculars, tools, and foul weather clothing.

## H. VOYAGE REPAIRS

This category includes the cost of unplanned repairs completed during the course of normal ship operation. This item was extracted from the category of maintenance and repair costs (which consists of planned maintenance, drydocking and voyage repairs). The remainder of the maintenance and repair items ( planned maintenance and drydocking) are included in the budgeted indirect cost category.

# IV. DATA ANALYSIS

A significant portion of the work behind the cost simulation tool lies in data analysis. Financial cost data is maintained on a mainframe computer located at Military Sealift Command (MSC) headquarters in Washington, DC. This detailed information includes the disposition of every payment voucher written by MSC. The information is accessible in read only form by MSCPAC and the other area commands via a Personal Computer (PC) modem hook up using software called the Financial Management Information System (FMIS) PC Gateway.

Data for this project was obtained from FMIS for all MSCPAC ships operated during fiscal years 1992 and 1993. The data consists of monthly summaries for the various cost categories required to operate ships. The total cost of operating a MSCPAC ship can be regarded as the sum of the following categories:

1. Salaries.
2. Training.
3. Fuel and lubricants.
4. Subsistence.
5. Port and miscellaneous (including spare parts).
6. Ship's equipage.
7. Maintenance and repair.
8. Budgeted overhead.

The monthly totals for each cost category were entered into a Microsoft Excel © spreadsheet database. This spreadsheet database allowed for ease in manipulating data to form variables to be used in the graphical data analysis software, A Graphical Statistical System © (AGSS). Additionally, Microsoft Excel is presently in use at MSCPAC, so every effort was made to stick to its use to reduce the amount of training needed for the future users of this application.

To ease the process of filtering the data to remove various undesirable data elements, Microsoft Excel's database query feature was used. For example, if it was desired to remove all data from months in which the ship was not available to the sponsor for 100 percent of the time, the query would be made requesting that those cases be deleted. The raw data can be similarly manipulated for any other desired case.

Graphical data analysis allows the characteristics of the data to be studied visually as well as through computational statistics. Sometimes it is far easier for the eye to see a relationship than it is to discern it from computational results. The cost category data can be modeled with known probability distributions and this relationship is easy to see graphically.

AGSS[1] is an interactive system for both two and three dimensional scientific-engineering graphics, applied statistics, and data

---

[1] IBM Corporation; <u>A Graphical Statistical System (AGSS) : An Introduction</u>; pg 1.

analysis. AGSS allows the user to create graphics, explore data interactively, analyze data using functions of applied statistics, develop customized graphics functions, and manage, review, and store work sessions. In this case, AGSS was chosen for its excellent capability in the area of distribution fitting.

AGSS has capabilities for fitting any of 18 univariate probability distributions to a set of data. The system computes maximum likelihood estimates as well as several other estimates of the distribution parameters, together with asymptotic confidence regions, and it can produce three dimensional contour and surface plots of the likelihood function.

AGSS provides a number of graphical comparisons of the empirical data with the theoretical fitted distribution. These displays help the user judge visually how well different distributional assumptions apply to the data. They provide graphs on which some representation of the theoretical distribution is superimposed on the corresponding empirical plot. Examples of this are plots of the histogram and fitted density function, empirical and fitted cumulative distribution functions (CDF), and probability plots. In a probability plot, if the empirical data corresponds exactly to the quantiles of the theoretical distribution, the points will lie exactly on the line that runs diagonally across the plot from lower left to upper right.

Kolmogorov-Smirnov (KS) bounds, at any desired confidence level, can be superimposed on the CDF and probability plots.

Four different goodness of fit tests are provided for uncensored data. These are the Chi-Square, Kolmogorov-Smirnov (KS), Cramer-Von Mises (C-VM), and Anderson-Darling (AD) tests. These tests give a quantitative measure of the goodness of fit. These values can be compared to the visual interpretation of the fit to help verify the quality of the fit. In graphical data analysis though, the visual impressions often give the analyst a far better indication of a probability distribution fit than do these statistical tests.

Figure 2 is a probability distribution plot created by AGSS. This view illustrates many of the principles discussed in the following paragraphs. Three plots are shown in this view. The plot in the upper left hand corner is a superimposed plot with a histogram of the empirical data superimposed on a plot of the theoretical probability density function selected for the fit. The plot in the upper right corner shows a plot of the empirical CDF superimposed on the theoretical CDF of the distribution selected for the fit. The plot in the lower left hand corner is a probability plot where the empirical percentiles are plotted against corresponding percentiles of the theoretical probability distribution. In this plot, if the empirical data exactly matched the theoretical probability distribution quantiles, all of the points

## T-AO 187 Class Salary

LOGISTIC CUMULATIVE DISTRIBUTION FUNCTION, N=25

LOGISTIC PROBABILITY PLOT, N=25
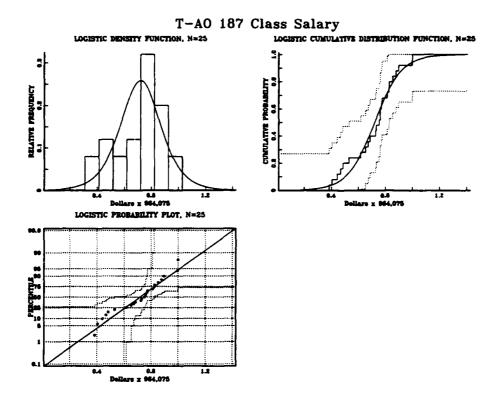
## Figure 2
## Sample Probability Distribution Plot

would plot on the straight line. A table of statistical and goodness of fit results normally is shown in the lower right hand corner of the plot. This four way plot is used extensively in the data analysis.

KS bounds are included on the CDF and probability plots. A theoretical CDF or straight line on the probability plot passing outside the bounds indicate lack of fit. The KS values are calculated for the user specified confidence level (95 percent in this analysis). The KS test statistic is based on the maximum difference between the observed empirical CDF and a hypothesized distribution across all values of $x$. As KS statistic values

24

decrease, the fit between the sample and the theoretical distribution is improved.

The significance level output by AGSS is the $p$-value used in statistics. The significance level is the smallest level of significance of the test statistic for which the null hypothesis would be rejected. In the cases examined in this project, the null hypothesis is that the data came from the distribution specified for the fit. The higher the significance level the more likely that the fit will be accepted.

The Cramér-Von Mises (C-VM) test statistic is based on the integral of the squared distance between the empirical and theoretical curves. The C-VM value, like the KS value, should be as low as possible. The C-VM significance level is given only in ranges for the C-VM statistic.

The Anderson-Darling (AD) test statistic is an attempt to overcome a drawback in both the KS and C-VM tests. Both the KS and the C-VM are not sensitive to departures from the null hypotheses that occur in the tails of the distribution[2]. This is improved in the Anderson-Darling test by using a weighted distance measure, the weight being the reciprocal of the standard deviation of the difference between the curve functions. The smaller the AD value, the more likely the fit is accepted.

[2] Lewis, Orav; Simulation Methodology for Statisticians, Operations Analysts, and Engineers: Volume 1; pg 369

Note that the Chi-Square test results are reported by AGSS for some of the probability distribution fit plots. The Chi-Square test is used when fitting discrete distributions or when continuous distributions can be modeled by grouping the data into mutually exclusive discrete bins such as in a histogram. The normally recognized lower limit for frequency in a bin for the Chi Square test is five. With the small size of the data sets (25 and 16 data points), this limit is tested in nearly every case, and the Chi-Square test will therefore not be used in this analysis.

The remainder of this chapter is organized as follows:

A. Assumptions - Description of the assumptions used in analyzing the MSCPAC data.

B. Summary of Analysis Results.

C. T-AO 187 Class Data Analysis - a detailed account of the data analysis of each cost category for ships of the T-AO 187 class.

D. T-ATF 166 Class Data Analysis - a detailed account of the data analysis of each cost category for ships of the T-ATF 166 class.

## A. ASSUMPTIONS.

For each data set (or cost category for a particular class of ships) numerous plots were made in an attempt to find which of several probability distributions best fit the data. Early in this process, it became clear that

26

most of the cost categories were subject to a fair degree of variability, depending on the operating schedule for the ship.

For example, a ship that is deployed performing services for the sponsor burns significantly more fuel than a ship of the same class that is in port frequently due to maintenance and administrative requirements. The same sort of variability holds in the other six categories of direct cost considered in the model (salaries, training, subsistence, port and miscellaneous costs, ship's equipage, and voyage repairs).

Fortunately, the MSCPAC Operations Department recently conducted a study[3] in which the operating tempos of MSCPAC ships were examined. In this study, it was decided to consider a ship in one of three states: available to sponsor, not available to sponsor due to maintenance requirements, and not available to sponsor due to administrative or training requirements. The state for each of the assigned ships was recorded for the last two fiscal years, and projected for the next five fiscal years.

When assigning a per-diem rate for the ships, one must assume that the ship is exclusively available to the sponsor for the period which is being paid for (i.e., ships available to sponsor 100 percent of the time he is paying for). After making this assumption, the data sets were scrutinized and only the months in which the ship is 100 percent available to the sponsor

---

[3] Military Sealift Command; MSCPAC Operations Department QMB, October 1993

(i.e., no days spent in satisfying maintenance or administrative requirements) were included in the data set. This filtering resulted in a significant reduction in variance of the observed costs.

Since the purpose of the model is to predict cost, it is reasonable to assume that in an operational status, no cost category should show a negative or zero total for a month. This assumption is important because the data set contains some instances in which zero and negative sums are carried as a matter of convenience for the accountants who later shift funds between accounts. This assumption permits filtering the data set to remove any months in which a negative or zero balance is shown for any category. The filtering removes some points that may bias the plots toward the artificial data.

Rather than including the entire maintenance and repair category in the direct cost portion of the model, only the maintenance and repair account for voyage repairs is included as a direct cost. This is done since the other accounts under maintenance and repair are not normally used in months where the ship is 100 percent available to the sponsor. Those accounts are maintenance and dry-docking costs which, for our purposes, are summed over the entire period and then divided to permit allocation of the costs for the period of time concerned in the model run. The voyage repair

account, on the contrary, takes into account the normal voyage repairs that are charged to the ship while it is in operation.

For this project, only the T-AO 187 class oilers and the T-ATF 166 class oceangoing tugs are analyzed. In both cases, there are a sufficient number of ships of these classes assigned to MSCPAC to comprise data sets of a large enough size that data analysis and the inferences drawn from the data analysis will apply.

Each ship class is analyzed separately. For these ship classes, the following cost categories are analyzed:

1. Salaries.
2. Training.
3. Fuel and lubricants.
4. Subsistence.
5. Port and miscellaneous.
6. Ship's equipage.
7. Voyage repairs.

## B. SUMMARY OF RESULTS

The theoretical distributions that were selected as the best approximations for each category are given in Table 1 for the T-AO 187 class and Table 2 for the T-ATF 166 class. The maximum likelihood estimates for the parameters, mean, standard deviation, and a subjective assessment of the quality of the fit are also given. In some cases it was difficult to choose

29

from a small set of distributions, and in other cases, none of the distributions seemed to fit very well. In every case, the distribution chosen is the one the author judged to best explain the data when all evidence was considered.

Both data sets were analyzed to determine if there was any correlation between cost categories. In only one pair of categories, T-ATF 166 Training versus Subsistence costs, is a significant correlation found. This correlation is incorporated into the Cost Simulation Tool model.

## Table 1.
### T-AO 187 CLASS DATA ANALYSIS RESULTS

| Cost Category | Distribution | Parameters | Mean | Std Dev | Fit Quality |
|---|---|---|---|---|---|
| Salary | Logistic | $\alpha$=692,601, $\beta$=95,974 | 692,601 | 174,179 | Good |
| Training | Weibull | C =1.5615, a =24,864 | 22,346 | 14,622 | Good |
| Fuel | Logistic | $\alpha$=202,026, $\beta$=53,822 | 202,026 | 97,624 | Fair |
| Subsistence | Logistic | $\alpha$=5,966, $\beta$=3,723 | 19,486 | 4,944 | Marginal |
| Port and Misc. | Weibull | C =2.7352, a =171,010 | 152,570 | 61,177 | Excellent |
| Ship's Equipage | Weibull | C =1.0616, a =21,037 | 20,414 | 18,890 | Good |
| Voyage Repairs | Gamma | $\alpha$=0.838, $\beta$=127,700 | 106,950 | 97,542 | Good |

## Table 2.
### T-ATF 166 CLASS DATA ANALYSIS RESULTS

| Cost Category | Distribution | Parameters | Mean | Std Dev | Fit Quality |
|---|---|---|---|---|---|
| Salary | Normal | $\mu$=123,880, $\sigma$=54,864 | 123,880 | 54,864 | Good |
| Training | Gamma | $\alpha$=1.3843, $\beta$=2,243 | 3,104 | 2,639 | Marginal |
| Fuel | Weibull | C=1.7563, a =49,847 | 44,409 | 26,105 | Good |
| Subsistence | Logistic | $\alpha$=3,433, $\beta$=795.7 | 3,670 | 1,894 | Marginal |
| Port and Misc. | Lognormal | $\mu$=10.075, $\sigma$=0.6869 | 30,047 | 23,333 | Marginal |
| Ship's Equipage | Lognormal | $\mu$=7.4952, $\sigma$=1.3561 | 4,513 | 10,379 | Good |
| Voyage Repairs | Lognormal | $\mu$=11.082, $\sigma$=1.068 | 115,010 | 167,790 | Good |

For a detailed coverage of the analysis of each cost category for each class of ship, the reader is encouraged to read the following sections of this chapter. Following the cost category analyses, the correlation analysis is found for each class of ship.

## C. T-AO 187 CLASS ANALYSIS

### 1. Salary Cost.

Figure 3 shows the probability distribution fit plot for T-AO 187 class salary cost. The plot shows a good fit using the logistic



T-AO 187 Class Salary
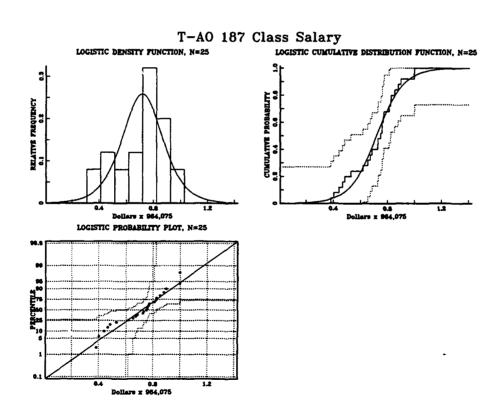
**Figure 3**
**T-AO 187 Class Salary Cost Probability Distribution Plot**

distribution. The theoretical CDF and distribution fall well within the chosen Kolmogorov-Smirnov (KS) bounds.

The goodness of fit tests indicate a good fit with KS value of 0.11668 and KS significance level of 0.88546. For the salary data, the C-VM value of 0.63255 and significance level > 0.15 indicate a good fit. For this plot, the AD value of 0.4681 and significance level of > 0.15 indicate that this is an adequate fit.

Although the fit looks good, there is some departure in the tails. This is seen in the plots of Figure 3. The values sharply start in the left tail and also sharply drop off in the right tail as seen in both the CDF view and the probability plot. Figure 4 is a view in which the empirical density has been superimposed over the theoretical density (empirical is solid curve, theoretical is dash-dot curve). In the superimposed plot, the departure in the tails is also seen as well as some difference in the area before the peak of the curve.

Figure 5 is a multiple box plot of the salary data. The actual data box closely resembles that of the small random sample box plot. The small random sample box plot was plotted from a random sample (using the fitted distribution and parameters) of the same size as the actual data set. The large random sample box plot was made from a random data set ten times the size of the actual data set. The large random sample box plot

T-AO 187 Salary
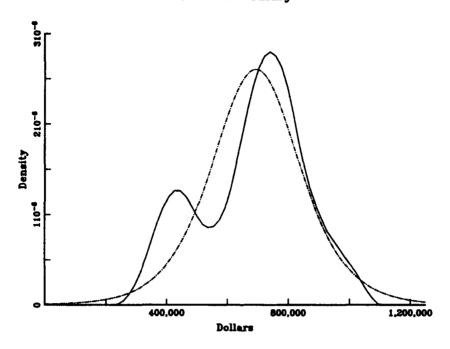


**Figure 4**
**T-AO 187 Class Salary Cost Superimposed Density Plot**

T-AO 187 Class Salary Cost
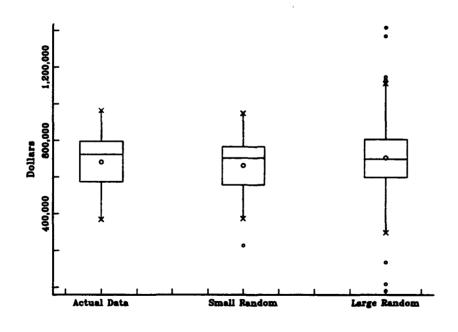


**Figure 5**
**T-AO 187 Class Salary Cost Multiple Box Plot**

33

closely resembles the actual data set plot except that the tails are longer due to the increased size of the data set.

The long lines extending above and below the actual data set box indicate some symmetry in the tails, although the lines are of slightly different length. The horizontal line across the box indicates the median and in this case since it is so far above the mean it shows some lack of symmetry near the peak of the curve. This is also seen in Figure 4. The data is skewed slightly to the left as indicated by the mean circle lying below the median line.

Figure 6 is a symmetry plot that agrees with the above analysis. The points lie below the $y=x$ line which indicates that the data is skewed to the left.

The fit with the logistic distribution is the best obtained for the salary cost data. The normal distribution is the closest runner up, but the goodness of fit results are not as strong. The fitted parameters for the logistic distribution are $692,601 and $95,974 for alpha and beta, respectively. The mean and standard deviation to be used in the Cost Simulation Tool Model are $692,601 and $174,179 respectively.

## 2. Training Cost.

Figure 7 is the probability distribution fit plot for monthly training cost data. The histogram seems to be almost tailor-made for the

34
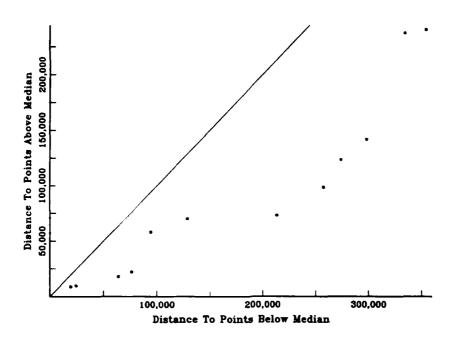
**Figure 6**
**T-AO 187 Class Salary Cost Symmetry Plot**
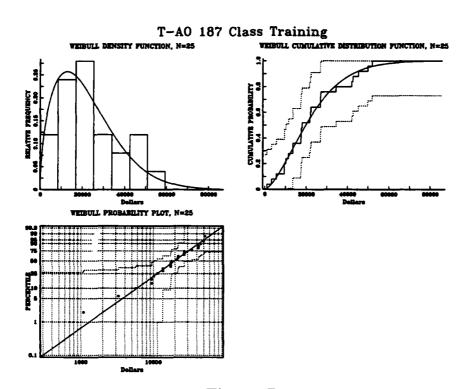


**Figure 7**

Weibull density function. The default bin selection by AGSS resulted in a histogram that is very similar to the shape of the density. Both the CDF and probability plots show the theoretical density to lie well within the KS bounds. The points in the probability plot of Figure 7 lie very close to the line.

The KS value of 0.09619 with significance of 0.97484 and the C-VM value of 0.031256 with significance > 0.15 indicate an excellent fit for the data with the Weibull distribution. The AD value of 0.24434 with significance > 0.15 indicates that the fit is good.

Figure 8 is the superimposed density plot showing the empirical data as a solid curve and the theoretical curve as a dot-dashed curve. Some lack of fit in the tails is seen in this view. The second peak seen in the empirical curve is due to one high leverage point in the data at approximately $45,000.

Figure 9 shows a multiple box plot for the data. The right skewness of the data set can be clearly seen in this view. This would be expected for data that were sampled from the Weibull distribution. The left tail is steeper and shorter than the right tail as indicated by the shorter line below the box. The median line is below the mean and shifted toward the bottom of the box. The small random sample box is slightly higher than the actual data set and large random sample data set boxes. Since the actual

T-AO 187 Training Costs
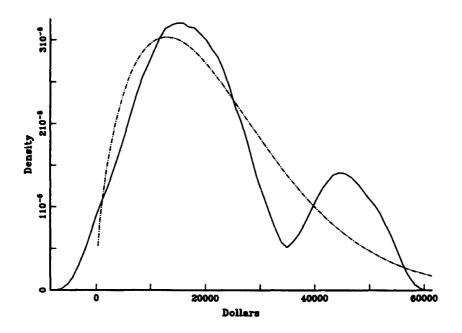


**Figure 8**
**T-AO 187 Class Training Cost Superimposed Plot**

T-AO 187 Class Training Cost
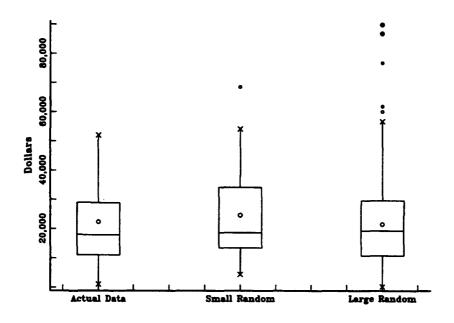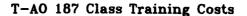


**Figure 9**
**T-AO 187 Class Training Cost Multiple Box Plot**

data set resembles the large random data set, the fit is good.

Figure 10 is a symmetry plot that also indicates a right skewness of the data. This is seen by the fact that the distances are much higher to points above the mean than to points below the mean.

The Weibull distribution will be used as the distribution for T-AO 187 class monthly training cost in the Cost Simulation Tool model. The gamma distribution is a close runner up, but, the Weibull fits are much better. The shape parameter to be used in the model will be 1.5615 and the
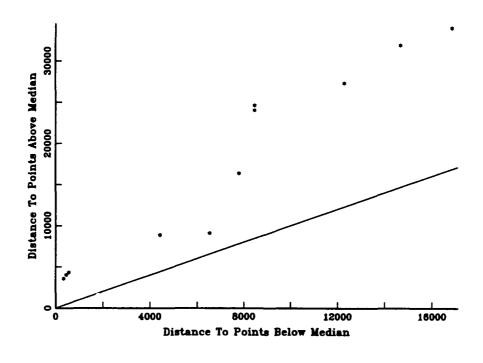


Figure 10
T-AO 187 Class Training Cost Symmetry Plot

scale parameter will be $24,864. The fitted mean and standard deviation for the data were $22,346 and $14,622, respectively.

### 3. Fuel Cost.

Figure 11 is a probability distribution fit plot for the fuel cost data. The best fit for the data is obtained with the logistic distribution. Although the superimposed density histogram plot shows what appears to be a good fit, it must be remembered that the appearance of the histogram can be significantly altered by changing the number of bins used. In this case, AGSS default values for bin selection provide a histogram that is a good match for the density. The CDF and probability plots indicate a lack of fit in the tails. The theoretical distribution is fully contained within the KS bounds for the data. The KS value is 0.15724 with a significance level of 0.56679. The C-VM value is 0.10111 with significance level > 0.15. The AD value of 0.71454 with significance level > 0.15 indicates an adequate fit.

Figure 12 is the superimposed density plot for the fuel cost data. As indicated above, the peak and middle area of the empirical density (solid curve) fit the theoretical density (dot-dashed curve) very well, but, the tails of the empirical density are thicker than those of the theoretical density.

Figure 13 is the multiple box plot for the fuel cost data. It shows that the majority of the empirical data set is tighter than either of the

## T-AO 187 Class Fuel Costs
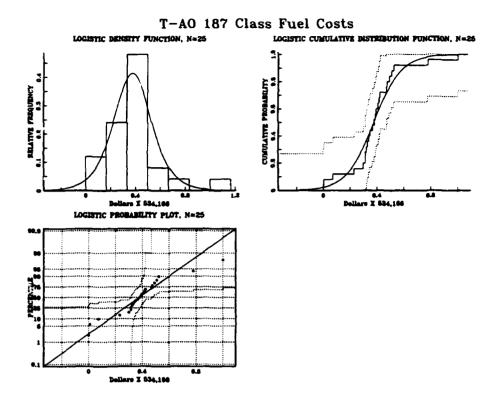


**Figure 11**
**T-AO 187 Class Fuel Cost Probability Distribution Plot**

random sample data sets. The view shows less similarity between the empirical data and the random samples than is indicated in Figures 11 or 12.

The spread of the actual data box is similar in both directions with outliers on both the left and right ends of the curve. Figure 14, the symmetry plot for the fuel cost data, also indicates strong symmetry with the points lying relatively even about the $y=x$ line.

The logistic distribution will be used in the Cost Simulation Tool model to represent fuel cost data. There were no other distributions that exhibited a good fit for the data. The fitted parameters are
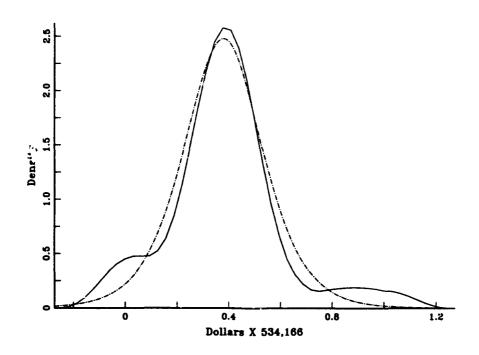
**Figure 12**
**T-AO 187 Class Fuel Cost Superimposed Density Plot**

$202,026 and $53,822 for alpha and beta, respectively. The fitted mean and standard deviation values are $202,026 and $97,624, respectively.

## 4. Subsistence Cost.

Figure 15 is the Probability Distribution Fit plot for subsistence cost. The superimposed histogram and density plot shows that the majority of the data lies in the two center bins. The CDF and probability plots show that the fit is less than perfect, but the theoretical distribution does lie within the KS bounds. The trace from the data departs from the theoretical shapes in the center due to the concentration of data there. The

41

T-AO 187 Class Fuel Cost
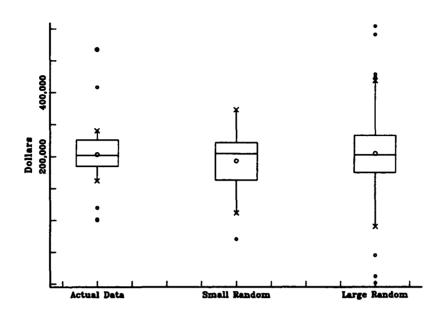


**Figure 13**
**T-AO 187 Class Fuel Cost Multiple Box Plot**

T-AO 187 Class Fuel Cost
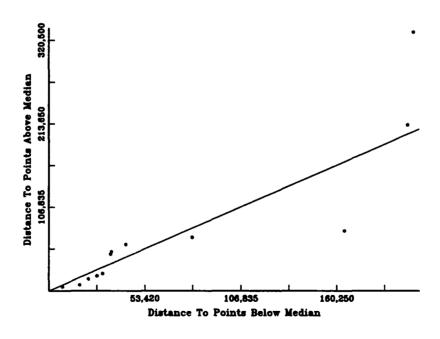


**Figure 14**
**T-AO 187 Class Fuel Cost Symmetry Plot**

42

## T-AO 187 Class Subsistence Cost

LOGISTIC DENSITY FUNCTION, N=25

LOGISTIC CUMULATIVE DISTRIBUTION FUNCTION, N=25

LOGISTIC PROBABILITY PLOT, N=25
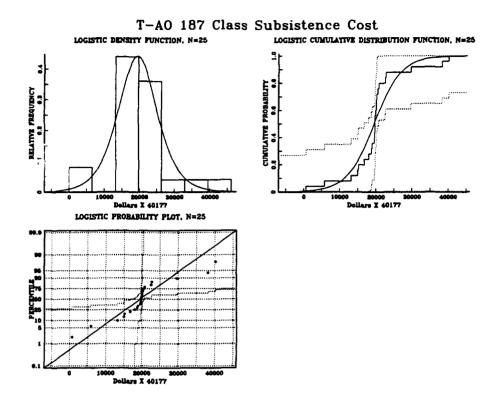
**Figure 15**
**T-AO 187 Subsistence Cost Probability Distribution Plot**

KS value of 0.21565 and significance level of 0.19535 indicates that there is a fit for the logistic distribution, albeit not a very strong one. Both the C-VM and AD suggest that there is an adequate fit, but the values for both tests, 0.27196 and 1.5123 respectively, are relatively large when compared with other values in this analysis. In both cases, the significance levels are > 0.15.

Figure 16, the superimposed density plot, indicates a much better fit than was evident in the previous plot, but the tails of the empirical density are much thicker than those of the theoretical density.

T-AO 187 Subsistence Cost
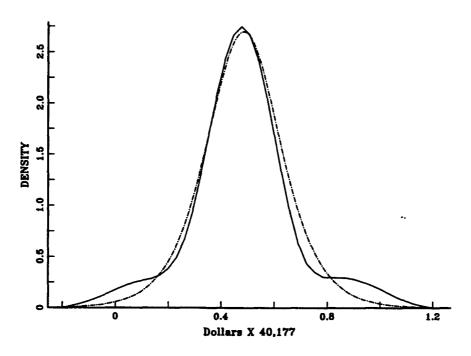


**Figure 16**
**T-AO 187 Subsistence Cost Superimposed Density Plot**

Even more of the probability is concentrated in the center of the empirical plot than is characteristic of the logistic distribution.

Figure 17 is the multiple box plot for the subsistence data. The actual data box does not remotely resemble the boxes for the random sample cases. The box and the extended lines of the actual data box are compressed and clearly indicate that the majority of the data is in the narrow peak of the distribution. There are five points recognized as distant outliers (the solid circles seen above and below the box).

44

**T-AO 187 Class Subsistence Cost**
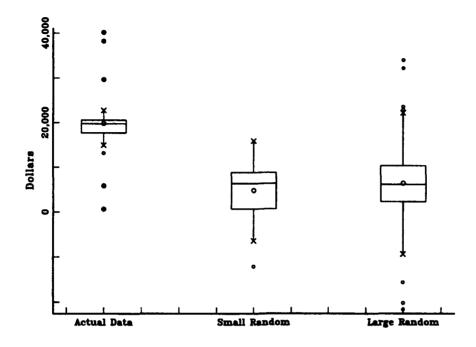


**Figure 17**
**T-AO 187 Class SubsistenceCost Multiple Box Plot**

Figure 18 is a symmetry plot for the data set. The plot indicates a symmetry which is consistent with the fit for the logistic distribution. This was also seen in the superimposed density plots.

The logistic distribution fit for this data is not very well supported, but it is the only fit obtainable from the 18 distributions available in AGSS. The logistic distribution will be used in the Cost Simulation Tool Model. The fitted parameters are $5,966 and $3,723 for alpha and beta. The

45
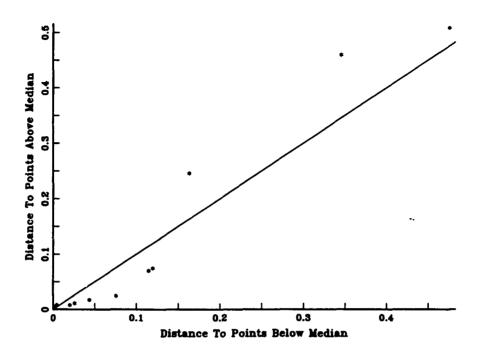
**Figure 18**
**T-AO 187 Class Subsistence Cost Symmetry Plot**

fitted mean and standard deviation were $19,486 and $4,944, respectively, and will be used as parameters in the model.

## 5. Port and Miscellaneous Cost.

Figure 19 is the Probability Distribution Fit plot for port and miscellaneous cost. The three plots all show a strong relationship between the data and the Weibull distribution. The CDF and probability plots show that the theoretical distribution is wholly included within the KS bounds. The stepped data traces very nearly mimic the CDF. The data

## T-AO 187 Class Port and Miscellaneous Cost

WEIBULL DENSITY FUNCTION, N=25

WEIBULL CUMULATIVE DISTRIBUTION FUNCTION, N=25

WEIBULL PROBABILITY PLOT, N=25
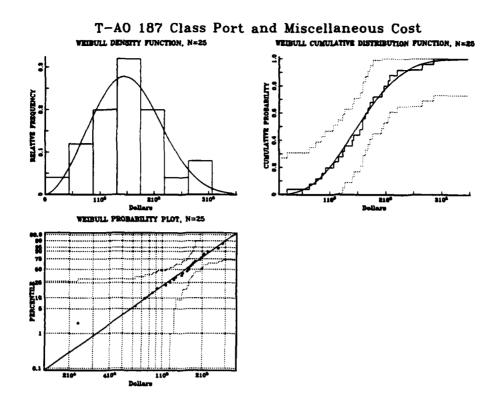
**Figure 19**
**T-AO 187 Port and Misc.Cost Probability Distribution Plot**

points all plot very close to the $y=x$ line of the probability plot. The KS value

of 0.089326 with significance level of 0.98844 indicate an excellent fit for the

Weibull distribution. The C-VM value of 0.026164 and significance level of

> 0.15 also indicate a strong fit. The AD value of 0.1888 with significance

level of >0.15 also indicate an adequate fit.

Figure 20 is the superimposed density plot for the data.

The empirical plot is very close to that of the theoretical density. A slight

lack of fit in the tails is visible in this view.
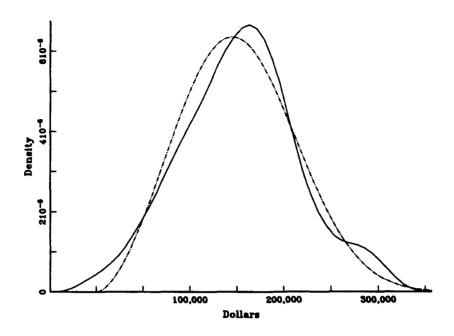
**Figure 20**
**T-AO 187 Port and Misc.Cost Superimposed Density Plot**

Figure 21, the multiple box plot for the port and miscellaneous cost, indicates a fairly symmetric distribution in the actual data. The actual data plot is similar to the large random sample plot, indicating a fit exists. The small random sample plot is slightly less symmetric and has a higher mean than the actual data plot. Figure 22 is the symmetry plot for the data. The points lie mostly below the $y=x$ line, but they are fairly close to the line which accounts for the symmetry.

The fit of the Weibull distribution to the data set is a strong one and will be used in the Cost Simulation Tool Model. The fitted shape and scale parameters are 2.7352 and $171,010, respectively. The

T-AO 187 Class Port and Misc Cost
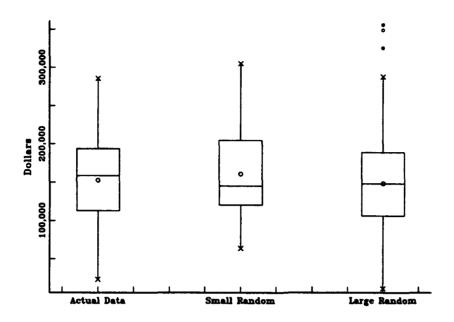


**Figure 21**
**T-AO 187 Port and Misc. Cost Multiple Box Plot**

T-AO 187 Class Port and Miscellaneous Cost
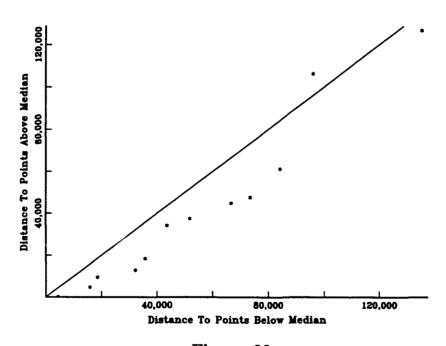


**Figure 22**
**T-AO 187 Port and Misc. Cost Symmetry Plot**

49

mean and standard deviation are $152,570 and $61,177, respectively.

## 6. Ship's Equipage Costs.

Figure 23 is the Probability Distribution Fit plot for ship's equipage cost. The data exhibits a good fit for the Weibull distribution. The CDF and probability plots clearly show that the theoretical distribution lies within the KS bounds. The stepped empirical CDF tracks fairly close to the theoretical CDF. The data points are close to the line in the probability plot.
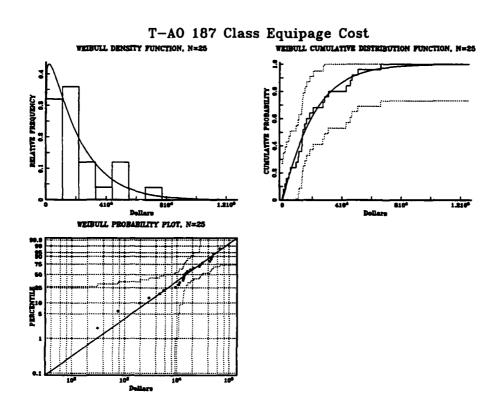


**Figure 23**
**T-AO 187 Ship's Equipage Cost Probability Distribution Plot**

The KS value of 0.10514 with a KS significance level of 0.94511 indicates a good fit exists. The C-VM value of 0.044199 with significance level > 0.15 also indicate a good fit. The AD results indicate an adequate fit exists with AD value of 0.29121 and significance level of > 0.15.

The superimposed density plot shown in Figure 24 displays less of a fit between the empirical density and the theoretical density than the test statistics would lead one to believe. There is clearly a lack of fit in both the left and right tails.

**T-AO 187 Equipage Cost**
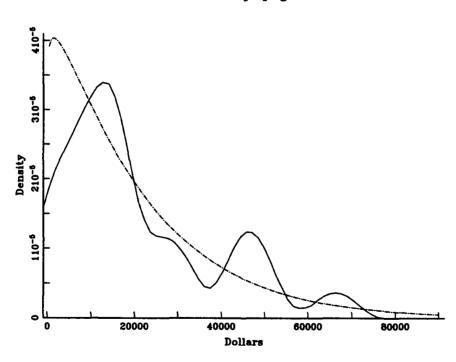


**Figure 24**
**T-AO 187 Ship's Equipage Cost Superimposed Density Plot**

51

Figure 25 is a multiple box plot for the ship's equipage cost data. The data appears to be skewed to the right as indicated by the shorter line below the box and the fact that the median is near the bottom of the box. The actual data plot closely resembles the large random sample plot. The small random sample plot is slightly different than the other two.

Figure 26 is a symmetry plot for the data set. It also supports the right skewness of the data with the long right tail indicated by the points lying above the line. The distance to points above the mean is much greater than the distance to points below the mean.

T-AO 187 Class Ships Equipage Cost
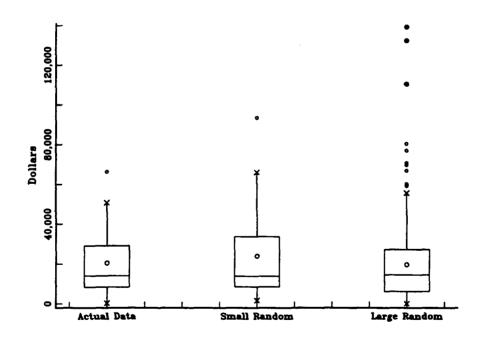


**Figure 25**
**T-AO 187 Ship's Equipage Cost Multiple Box Plot**
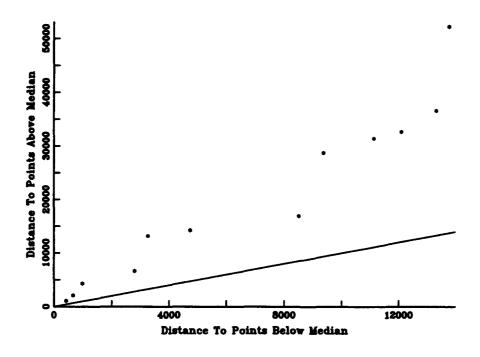
**Figure 26**
**T-AO 187 Ship's Equipage Cost Symmetry Plot**

  The fit for the Weibull distribution will be used for this category in the Cost Simulation Tool Model. There is also a good fit for the gamma distribution with this data, but the fit is not as strong. The fitted shape and scale parameters of 1.0616 and 21,037, respectively, for the Weibull distribution will be used in the model. The mean and standard deviation for the data were $20,414 and $18,890, respectively.

## 7. Voyage Repair Cost.

Figure 27 is the probability distribution fit plot for the voyage repair cost data. The superimposed histogram and density plot is not of much use in this case due to lack of detail in the default view presented by AGSS. The relative frequency scale in this view seems to be too large for the data presented by the histograms. A good fit with the gamma distribution is indicated by the information in this figure. The CDF and probability plots both show that the theoretical distribution is contained within the KS bounds. The KS value of 0.11317 with significance level of 0.90803 is
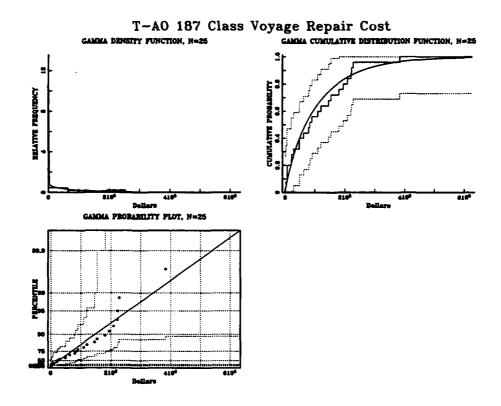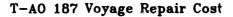


**Figure 27**
**T-AO 187 Voyage Repair Cost Probability Distribution Plot**

another indicator of a good fit. The C-VM value of 0.072187 and significance level > 0.15 also indicates a fair fit. The AD indicates an adequate fit with a value of 0.47551 and significance level > 0.15 respectively.

The superimposed density plot of Figure 28 shows less of a resemblance between the empirical density and the theoretical density than the test results above seem to indicate. The 'ack of fit in the tails is also clearly seen in this view. The data also is skewed to the right in this view.

Figure 29 is a multiple box plot of the voyage repair cost data. The right skewness of the actual data is very pronounced in this view.

T-AO 187 Voyage Repair Cost
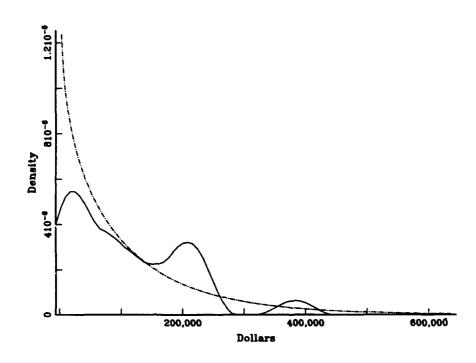


**Figure 28**
**T-AO 187 Voyage Repair Cost Superimposed Density Plot**

T-AO 187 Class Voyage Repair Cost
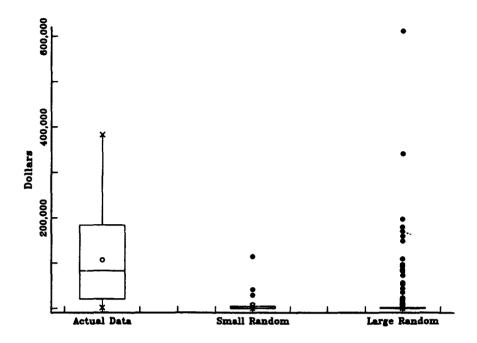


**Figure 29**
**T-AO 187 Voyage Repair Cost Multiple Box Plot**

The left tail is short as evidenced by the short line extending from the bottom of the box. The right tail, on the contrary, is very long as seen by the long line extending from the top of the box. The median line is toward the bottom of the box and lies below the mean. The random sample views are not of much help in this case. Based on this plot, an inference that a fit does not exist would be drawn.

Figure 30 is a symmetry plot that also shows the long right tail by the difference in the axes and the fact that the points plot above the $y=x$ line and depart further from the line as the cost increases.

T-AO 187 Class Voyage Repair Cost
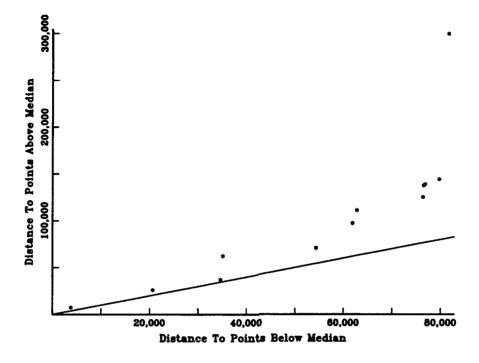


**Figure 30**
**T-AO 187 Voyage Repair Cost Symmetry Plot**

The gamma distribution fit will be used for voyage repair cost in the Cost Simulation Tool Model. The runner up is a fit with the lognormal distribution that is not as strong as the fit with the gamma distribution. The fitted parameters for the gamma distribution are 0.83755 for alpha and 127,700 for beta. The mean and standard deviation are $106,950 and $97,542, respectively.

## 8. Correlation Between Data Categories.

The multiple correlation test feature in AGSS is used to test for correlation between any pair of the categories of data. The presence

57

of any correlation would be important to explain during the cost analysis

phase of the project. If correlations exist, they will also require

compensation in the Cost Simulation Tool Model. Crystal Ball, the

simulation add-in software for Microsoft Excel, allows the user to input

Spearman Rank correlation between any tw          ; used in a model.

The multiple correlation test for the T-AO 187 data

matrix indicates that no significant correlation exists between any pair of

data sets in the data matrix. Table 3 is the multiple correlation matrix for

the T-AO 187 Class data.

**Table 3.**
**T-AO 187 CLASS DATA MULTIPLE CORRELATION MATRIX**

|               | Salaries | Training | Fuel   | Subsistence | Port   | Equipage | Voyage Repairs |
|---------------|----------|----------|--------|-------------|--------|----------|----------------|
| Salaries      | 1        | 0.0694   | 0.191  | 0.246       | 0.339  | -0.251   | 0.137          |
| Training      | 0.0694   | 1        | -0.177 | -0.121      | 0.055  | -0.228   | 0.138          |
| Fuel          | 0.191    | -0.177   | 1      | -0.118      | 0.142  | -0.172   | -0.13          |
| Subsistence   | 0.246    | -0.121   | -0.118 | 1           | -0.121 | -0.056   | 0.096          |
| Port          | 0.339    | 0.055    | 0.142  | -0.121      | 1      | -0.149   | 0.146          |
| Equipage      | -0.251   | -0.228   | -0.172 | -0.056      | -0.149 | 1        | 0.153          |
| Voyage Repairs| 0.137    | 0.138    | -0.13  | 0.096       | 0.146  | 0.153    | 1              |

## D. T-ATF 166 CLASS ANALYSIS.

### 1. Salary Cost.

Figure 31 is the probability distribution fit plot for the

T-ATF 166 class salary cost data. A normal distribution fit was obtained for

T-ATF 166 Class Salary Cost

NORMAL DENSITY FUNCTION, N=16

NORMAL CUMULATIVE DISTRIBUTION FUNCTION, N=16

NORMAL PROBABILITY PLOT, N=16
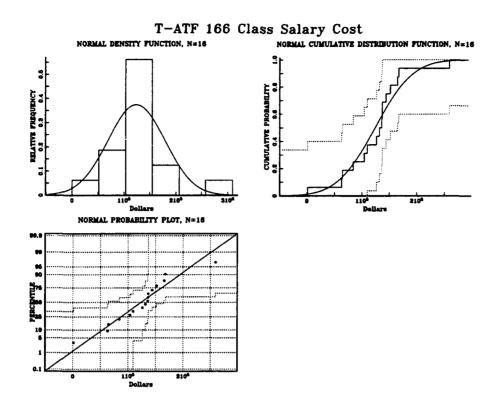
**Figure 31**

**T-ATF 166 Class Salary Cost Probability Distribution Plot**

this data set. The superimposed histogram-density plot shows a wide dispersion of values with an exceptionally large density in the center bin of the histogram. The CDF and probability plots show that the theoretical distribution is completely contained within the KS bounds. The KS value of 0.15001 and significance level of 0.86422 indicate a fair fit between the empirical data and the theoretical density. The C-VM results of 0.075901 and significance > 0.15 also support this hypothesis. The AD test indicates a fit exists with a value of 0.47923 and significance > 0.15.

Figure 32 is the superimposed density plot for the salary
cost data. As mentioned above, the plot shows that the fit with the normal
distribution is generally good, but that there is a lack of fit in the tails of the
distribution.

Figure 33 is the multiple box plot for the data. This view
indicates that the actual data is similar to the large random sample, but
slightly different in spread from the small random sample box.

Figure 34 is the symmetry plot for the salary cost data. A
right skewness is indicated by the information in this plot. The distances to

**T—ATF 166 Class Salary Cost**
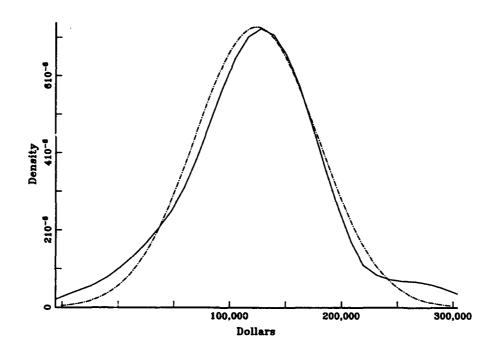


**Figure 32**
**T-ATF 166 Class Salary Cost Superimposed Density Plot**
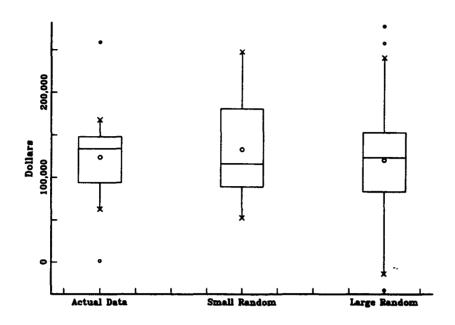
**Figure 33**
T-ATF 166 Class Salary Cost Multiple Box Plot

T-ATF 166 Class Salary Cost
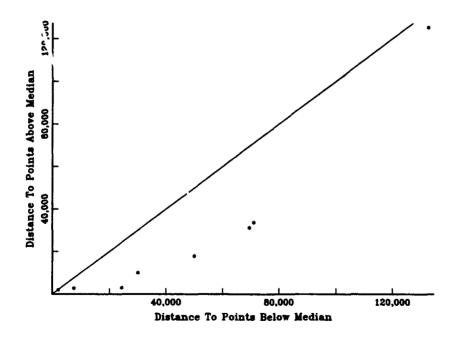


**Figure 34**
T-ATF 166 Class Salary Cost Symmetry Plot

61

points below the median are generally greater than the distances to points above the median.

The normal distribution fit will be used in the Cost Simulation Tool model. There are no other distributions that are considered good fits for this data set. The fitted mean and standard deviation are $123,880 and $54,864, respectively.

## 2. Training Cost.

Figure 35 is the probability distribution fit plot for the training cost data. An acceptable fit was obtained using the gamma distribution. The superimposed histogram-density plot shows that th histogram has the same general shape as the gamma density. The CDF and probability plots show that the theoretical distribution is contained within the KS bounds. The stepped empirical CDF does not adhere closely to the theoretical CDF in some areas of the plot. The same is true in the probability plot where the data points cross back and forth over the line in the right tail of the plot. The KS value of 0.20696 with significance level of 0.4996 does not give as strong a case for a fit as some of the other data sets seen earlier. The C-VM value of 0.08997 with significance level >0.15 supports the hypothesis that a fit exists. The AD value of 0.4847 with significance level of > 0.15 indicates that a fit exists.

**T-ATF 166 Class Training Cost**

GAMMA DENSITY FUNCTION, N=16     GAMMA CUMULATIVE DISTRIBUTION FUNCTION, N=16

GAMMA PROBABILITY PLOT, N=16

**Figure 35**
**T-ATF 166 Class Training Cost Probability Distribution Plot**

Figure 36 is the superimposed density plot for the data. In this view, the left part of the data set up to the peak seems to be the area where there is a departure from the theoretical gamma density.

Figure 37 is a multiple box plot which shows the data to be right skewed. This agrees with what was seen in the superimposed density plot. The data would be left skewed if not for the two largest values in the data set. This is shown by both of the random sample views.

Figure 38 is a symmetry plot that also shows the strong right skewness of the data. The distance to most of the points above the

T-ATF 166 Class Training Cost
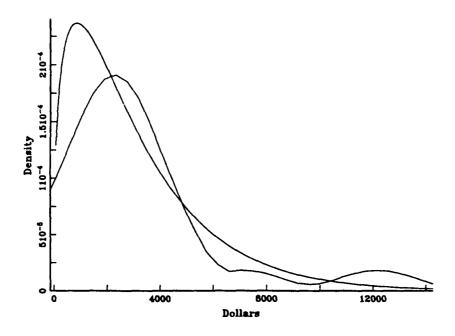


**Figure 36**
**T-ATF 166 Class Training Cost Superimposed Density Plot**

T-ATF 166 Class Training Cost
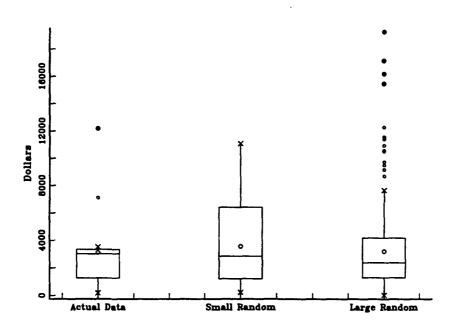


**Figure 37**
**T-ATF 166 Class Training Cost Multiple Box Plot**

64

T-ATF 166 Class Training Cost
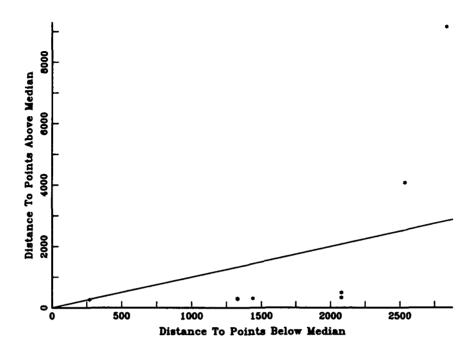


**Figure 38**
**T-ATF 166 Class Training Cost Symmetry Plot**

median is less than $1000, whereas the distances to points below the median ranges from $1250 to $2500.

The gamma distribution fit will be used in the Cost Simulation Tool model for T-ATF 166 monthly training cost. A fit is also obtained using the Weibull distribution, but the fit is not as good as this fit with the gamma distribution. The fitted parameters for the gamma distribution are 1.3843 for alpha and 2,242.50 for beta. The mean and standard deviation were $3,104.40 and $2,638.50, respectively.

65

### 3. Fuel Cost.

Figure 39 is the probability distribution fit plot for the fuel cost data. These views indicate a good fit of the fuel cost data to the Weibull distribution. The superimposed histogram-density plot shows that there is a good resemblance between the histogram and the theoretical Weibull distribution. The stepped empirical CDF tracks fairly closely to the theoretical CDF. The data points plot close to the line in the probability plot. The theoretical distribution is contained within the KS bounds in both the CDF and probability plots. The KS value of 0.13903 and significance of 0.91656 support the hypothesis that a fit exists. The C-VM value of 0.032114 and significance level of > 0.15 also support the hypothesis. The AD value of 0.211 with significance level of > 0.15 indicate a fit exists.

Figure 40 is the superimposed density plot. The shape of the empirical density is very close to that of the theoretical density with some small departures near the peak and in the right tail. Both the empirical and theoretical densities are skewed slightly to the right.

Figure 41 is the multiple box plot of the data. There is similarity between the actual data box and the random sample boxes. This view agrees with the slight right skewness seen in the superimposed density plot earlier. The line extending from the bottom of the box is smaller than

## T-ATF 166 Class Fuel Cost
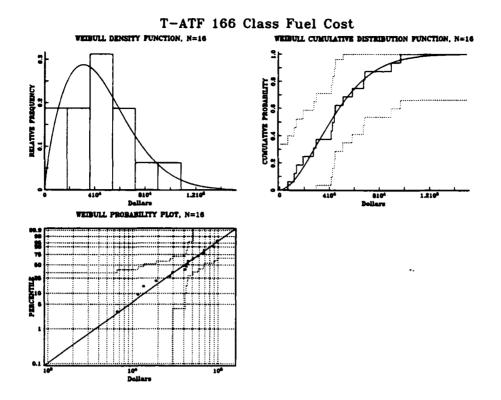


**Figure 39**
**T-ATF 166 Class Fuel Cost Probability Distribution Plot**

the line extending above the box indicating a longer right tail. The median is near the middle of the box and is very close to the mean value.

Figure 42 is the symmetry plot of the data. The points below the median are all closer to it than the points above the median. This also agrees with the slight right skewness and long right tail shape discussed above.

The Weibull distribution fit will be used for fuel cost data in the Cost Simulation Tool Model. An acceptable fit is also obtained using the gamma distribution, but the fit is not as good as the Weibull distribution
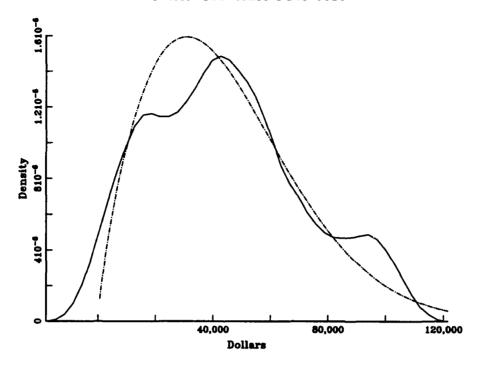
**Figure 40**
**T-ATF 166 Class Fuel Cost Superimposed Density Plot**

fit. The fitted parameters for the Weibull distribution were 1.7563 for shape

and 49,847 for scale. The mean and standard deviation were $44,409 and

$26,105, respectively.

## 4. Subsistence Cost.

Figure 43 is the probability distribution fit plot for the

subsistence cost data. The plots indicate that a fair fit exists for the data

with the logistic distribution. The superimposed histogram-density plot

shows that the histogram shape is similar to that of the density. The stepped

empirical CDF departs somewhat from the theoretical CDF due to the high

68
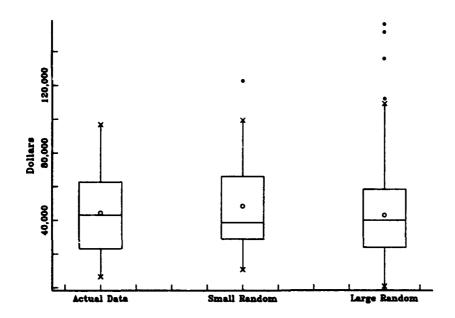
**Figure 41**
**T-ATF 166 Class Fuel Cost Multiple Box Plot**

T-ATF 166 Class Fuel Cost
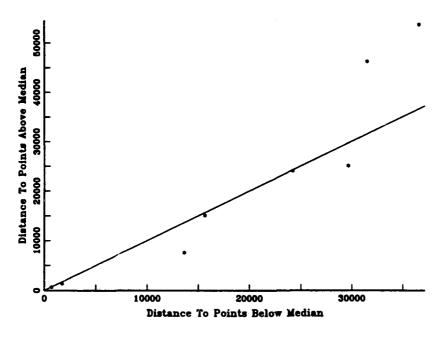


**Figure 42**
**T-ATF 166 Class Fuel Cost Symmetry Plot**

69
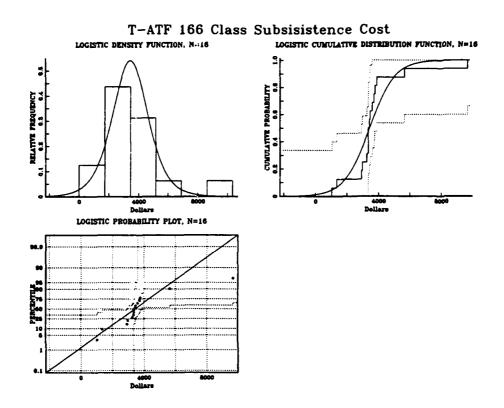
**Figure 43**
**T-ATF 166 Class Subsistence Cost Probability Distribution Plot**

density of data in the center, but the theoretical CDF is completely contained within the KS bounds. In the probability plot, the points are not very close to the line although the line is completely contained within the area of the KS bounds. The KS value of 0.23775 with significance of 0.32626 does support the hypothesis that a fit exists. The C-VM value of 0.2226 and significance level > 0.15 do not rule out a fit. The AD value of 1.2804 with significance level > 0.15 indicate that a fit exists.

Figure 44 is the superimposed density plot for the subsistence cost data. This view presents a much stronger case for using the

70
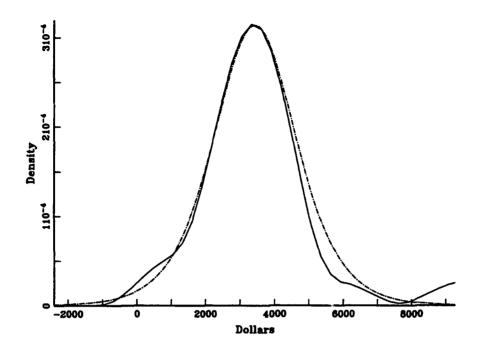
**Figure 44**
**T-ATF 166 Class Subsistence Cost Superimposed Density Plot**

logistic distribution fit. The empirical density very nearly duplicates the shape of the theoretical density. There is some slight departure in the tails of the distribution.

Figure 45 is the multiple box plot for the data. This plot shows that symmetry exists for the data set. The actual data is more tightly squeezed in the middle of the distribution than the random sample boxes. The tight grouping of the data does conform to the tight peaked appearance of the logistic distribution though it is slightly tighter which accounts for the crossing back and forth over the line in the probability plot.
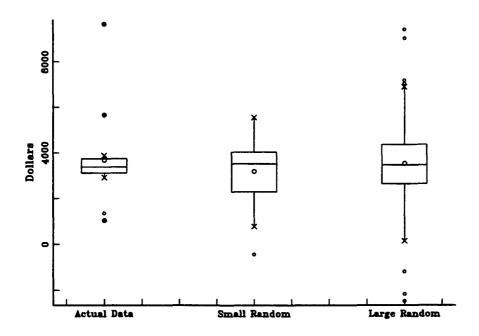
**Figure 45**
**T-ATF 166 Class Subsistence Cost Multiple Box Plot**

Figure 46 is the symmetry plot for the data. This plot also indicates symmetry for the data with only one data point departing from the vicinity of the $y=x$ line.

The logistic distribution will be used for the T-ATF 166 subsistence cost in the Cost Simulation Tool model. There are no other distributions that exhibited a good fit for this data. The fitted parameters are 3432.7 for alpha and 795.66 for beta. The mean and standard deviation are $3669.80 and $1894, respectively.
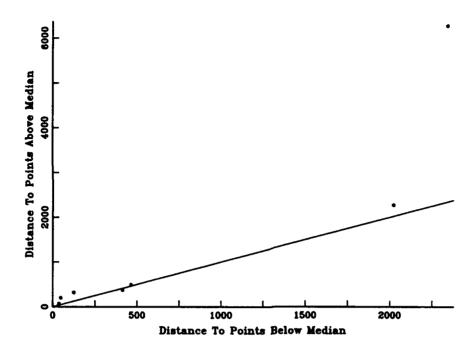
**Figure 46**
**T-ATF 166 Class Subsistence Cost Symmetry Plot**

## 5. Port and Miscellaneous Cost.

Figure 47 is the probability distribution fit plot for port and miscellaneous cost. From the analysis in this plot there is a marginal fit to the lognormal distribution. Most of the data is grouped tightly about $20,000, but there is one value of over $400,000 and two values in the $100,000 range that exert leverage to shift the density toward the extreme right. A satisfactory fit with any distribution was unattainable until the point at $400,000 was removed. A marginal lognormal fit was then obtained. With the aforementioned caveat, the theoretical distribution in the CDF and

73

T-ATF 166 Class Port and Miscellaneous Cost
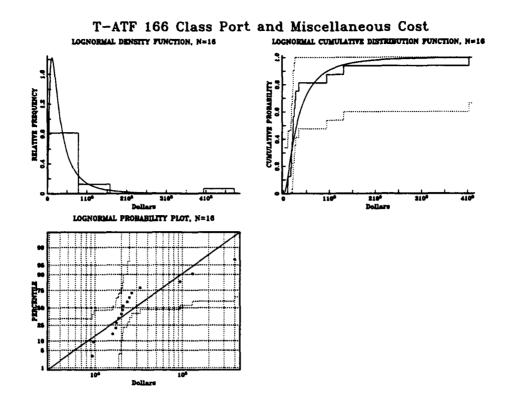
**Figure 47**
**T-ATF 166 Port and Misc. Cost Probability Distribution Plot**

probability plots falls entirely within the KS bounds. The KS value of 0.239932 with significance of 0.35671 satisfy the hypothesis that a fit exists. The C-VM value of 0.18118 with significance level > 0.15 does not rule out the hypothesis that a fit exists. The AD value of 1.0456 with significance level > 0.15 also indicate a fit.

The superimposed density plot shown in Figure 48 indicates a lack of fit where there is a gap in the empirical data between the clump of values at about $20,000 and the large values mentioned previously.

74
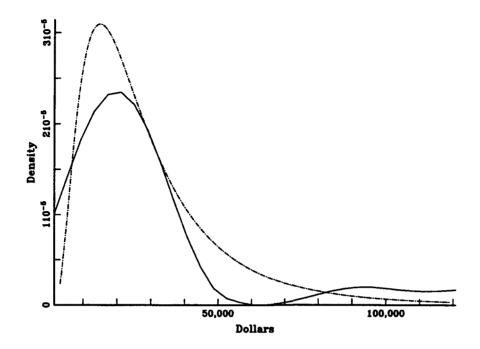
**Figure 48**
**T-ATF 166 Port and Misc. Cost Superimposed Density Plot**

Figure 49 is a multiple box plot of the port and miscellaneous cost data. The distribution is mostly a tight bunch of values as can be seen by the tightly squeezed box at the lower values. There are also 4 outlier points which account for the long right tail of the data. The actual data box is similar to the large random sample box.

Figure 50 is a symmetry plot that shows that the points located in the tight bunch are symmetric, but points in the right tail cause a radical departure from the symmetry.
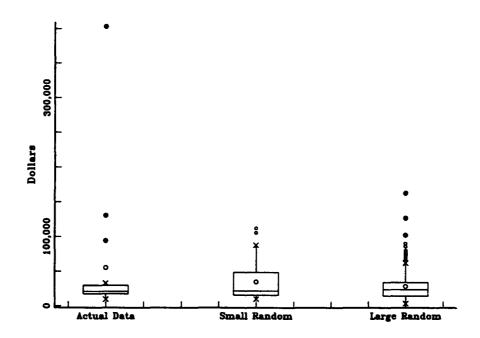
**Figure 49**
**T-ATF 166 Port and Misc. Cost Multiple Box Plot**

The only fit obtainable, as mentioned before, is achieved after removing the point at $400,000. Discussion with both the MSCPAC accounting and operations departments indicates that $400,000 for one month's port and miscellaneous cost is much greater than normal and would not be expected to be repeated. The lognormal distribution will be used in the Cost Simulation Tool model. The fitted parameter values are 10.075 and 0.68694 for mu and sigma, respectively. The values for the mean and standard deviation are $30,047 and $23,333, respectively.
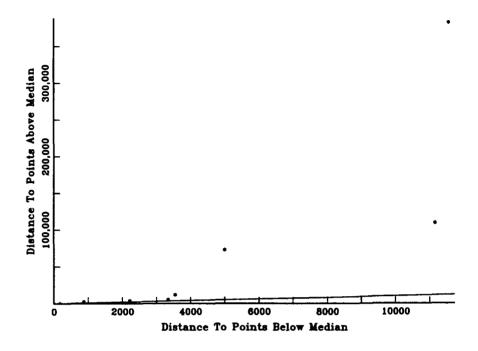
**Figure 50**
**T-ATF 166 Port and Misc. Cost Symmetry Plot**

## 6. Ship's Equipage Cost.

Figure 51 is the probability distribution fit plot for the ship's equipage data. This data was found to satisfy a fit for the lognormal distribution. The superimposed histogram-density plot shows a good resemblance between the histogram and the theoretical density. The stepped empirical CDF tracks fairly well with the theoretical CDF. The data points fall close to the line in the probability plot. The theoretical distributions lie completely within the KS bounds in both plots. The KS value of 0.13769 with a significance level of 0.92204 indicates that a fair fit exists. The C-VM
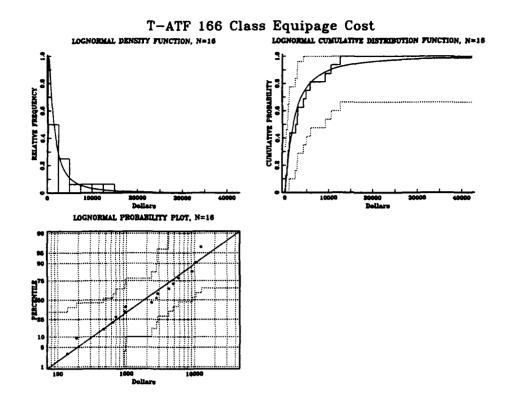
**Figure 51**
**T-ATF 166 Ship's Equipage Cost Probability Distribution Plot**

value of 0.050035 with significance level > 0.15 also supports the hypothesis

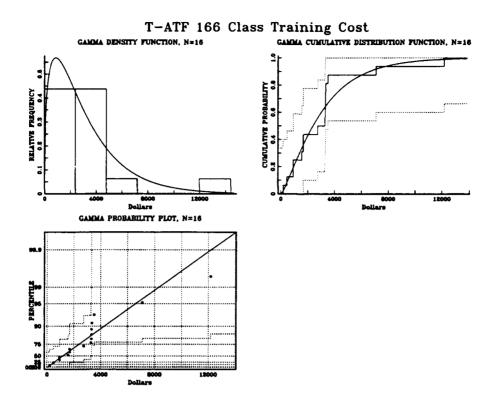that a fit exists. The AD result of 0.31448 with significance > 0.15 indicates a

fit.

Figure 52 is the superimposed density plot which shows that

there is some similarity in the shape of the empirical density and the

theoretical density. There are a few data points in the $10,000 range that

cause a lack of fit in the right tail.

Figure 53 is the multiple box plot for the ship's equipage data.

This plot shows that the data are right skewed with a shorter line below the
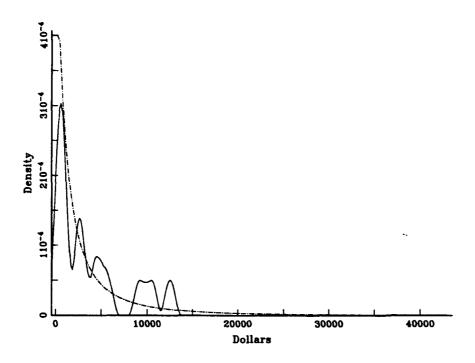
**Figure 52**
**T-ATF 166 Ship's Equipage Cost Superimposed Density Plot**

box and the median below the middle of the box as well as being below the mean value. The right tail is much longer than the left tail primarily due to the data points in the $10,000 range. The actual data box is similar to the large random sample box plot.

Figure 54 is the symmetry plot of the data. This view also supports the idea that the data are right skewed. Most of the distances to points above the mean are longer than distances to points below the mean.

The lognormal distribution will be used to model ship's equipage cost in the Cost Simulation Tool model. The Weibull distribution also fit the
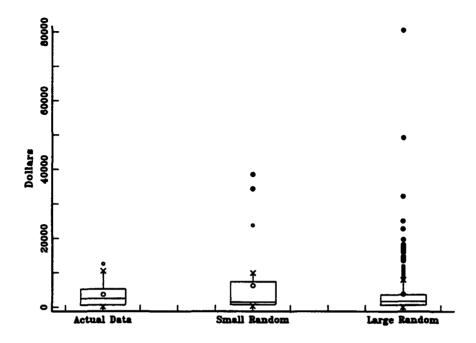
**Figure 53**
**T-ATF 166 Ship's Equipage Cost Multiple Box Plot**

data, but the fit with the lognormal distribution is much better. The fitted

parameters for the lognormal distribution are 7.4952 and 1.3561 for mu and

sigma, respectively. The fitted mean and standard deviation are $4512.80

and $10,379 respectively. The large standard deviation is cause for concern

here. The addition of more data points after more operation time will help in

re-plotting, fitting, and finding a fit with reduced variance.

T-ATF 166 Class Equipage Cost
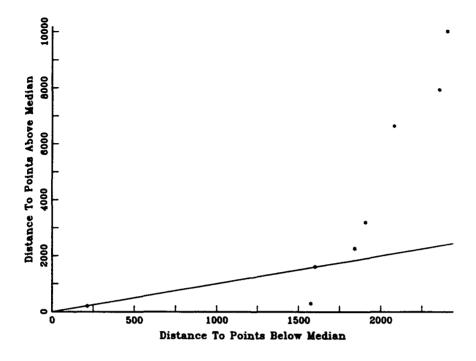


**Figure 54**
**T-ATF 166 Ship's Equipage Cost Symmetry Plot**

## 7. Voyage Repair Cost.

Figure 55 is the probability distribution fit plot for the voyage repair cost data. A fit was obtained with the lognormal distribution. The superimposed histogram-density plot shows that although there is some resemblance between the histogram and the data, there are also gaps in the data at several places. The stepped empirical CDF does exhibit a fit with the theoretical CDF. The data points plot near the line in the probability plot. The theoretical distribution in both plots lies within the KS bounds. The KS value of 0.12223 with significance of 0.9706 indicates that a good fit exists.

81

## T–ATF 166 Class Voyage Repair Cost

LOGNORMAL DENSITY FUNCTION, N=16    LOGNORMAL CUMULATIVE DISTRIBUTION FUNCTION, N=16

LOGNORMAL PROBABILITY PLOT, N=16
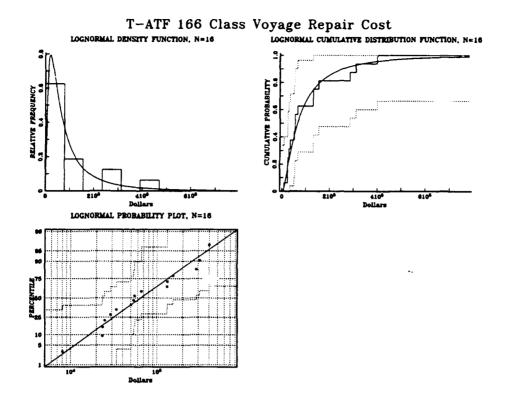
**Figure 55**
**T-ATF 166 Voyage Repair Cost Probability Distribution Plot**

The C-VM value of 0.052749 with significance level > 0.15 supports the fit.

The AD value of 0.3337 with significance level > 0.15 indicates a fit exists.

Figure 56 is the superimposed density plot for the voyage repair cost data. The empirical density seems to have a good match on the left side of the curve, but on the right side the data points with large gaps in between cause a departure from the natural shape of the theoretical distribution.

Figure 57 is a multiple box plot of the data. The right skewed nature of the data is clearly seen in this view where the line below
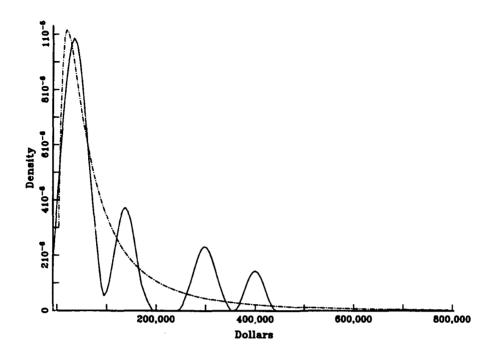
**Figure 56**
**T-ATF 166 Voyage Repair Cost Superimposed Density Plot**

the box is significantly shorter than the one above the box. This indicates

that the right tail is much longer than the left tail. The median line is well

below the middle of the box and lies below the mean as well. The actual data

box resembles the small random sample except that the actual data box is

slightly narrower.

Figure 58 is a symmetry plot of the data. This view

clearly shows the right skewness of the data. The distance to points above

the mean is much greater than the distance to points below the mean. All of

the points plot well above the $y=x$ line.

T-ATF 166 Ships Voyage Repair Cost
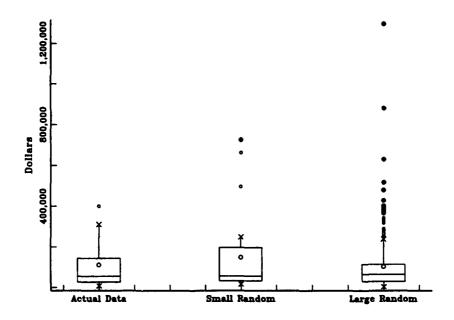


**Figure 57**
**T-ATF 166 Voyage Repair Cost Multiple Box Plot**

T-ATF 166 Class Voyage Repair Cost
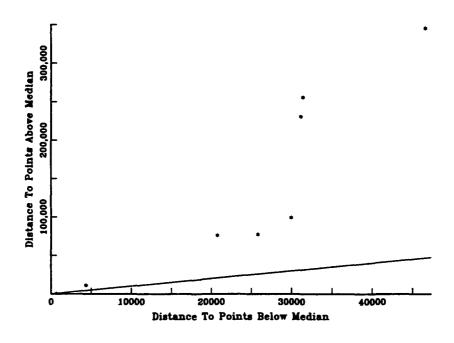


**Figure 58**
**T-ATF 166 Voyage Repair Cost Symmetry Plot**

84

The lognormal fit for this data set will be used in the Cost Simulation Tool model. The data also fit the Weibull distribution, but the fit was not as good. The fitted parameters for the lognormal distribution are 11.082 and 1.068 for mu and sigma. The mean and standard deviation are $115,010 and $167,790, respectively.

### 8. Multiple Correlation Test Results.

The T-ATF 166 class data matrix was examined using the multiple correlation function in AGSS to determine if there were any significant correlations between data set pairs. Table 4 is the pairwise correlation matrix.

From the matrix in Table 4, there are three cases of possibly significant correlation. Correlation seems to exist between salary cost and voyage repair cost (0.396), training cost and subsistence cost (0.372), and port and miscellaneous cost and voyage repair cost (0.639). Each of these correlations is further examined in a pairwise manner. In each case, a bivariate scatter plot is plotted. Then linear regression is performed to permit further graphical examination of the relationship between the data sets. In all three cases, there is one high leverage point which seems to be driving the correlation. In each case, the high leverage point is removed, the data is replotted, and the correlation is recalculated.

**Table 4**
**T-ATF 166 CLASS MULTIPLE CORRELATION DATA MATRIX**

|  | Salary | Training | Fuel | Subsistence | Ports & Misc | Equipage | Voyage Repairs |
|---|---|---|---|---|---|---|---|
| Salary | 1 | 0.091 | -0.05 | -0.152 | 0.094 | 0.0388 | 0.396 |
| Training | 0.091 | 1 | 0.108 | 0.372 | -0.204 | -0.004 | -0.045 |
| Fuel | -0.05 | 0.108 | 1 | 0.151 | 0.193 | 0.173 | 0.113 |
| Subsistence | -0.15 | 0.372 | 0.151 | 1 | 0.213 | -0.317 | -0.09 |
| Ports & Misc | 0.094 | -0.204 | 0.193 | 0.213 | 1 | 0.173 | 0.639 |
| Equipage | 0.388 | -0.004 | 0.173 | -0.317 | 0.173 | 1 | 0.249 |
| Voyage Repairs | 0.396 | -0.045 | 0.113 | -0.09 | 0.639 | 0.129 | 1 |

Figure 59 is the plot of salary versus voyage repairs. After the high leverage point is removed, the correlation falls to such a low level that it is considered insignificant.

Figure 60 is the plot of training versus subsistence cost after removal of the high leverage point. With the high leverage point removed, the correlation actually increases. In this case, a Spearman Rank Correlation of 0.894 will be used in the Cost Simulation Tool Model. During the cost analysis phase of the project, an explanation for this high correlation will be sought.

Figure 61 is the plot of port and miscellaneous cost versus voyage repairs. With the high leverage point removed, the correlation is reduced to the point where it will be considered insignificant.

.

ATF SALARY VS VOYAGE REPAIRS (LESS HI LEVERAGE PT)
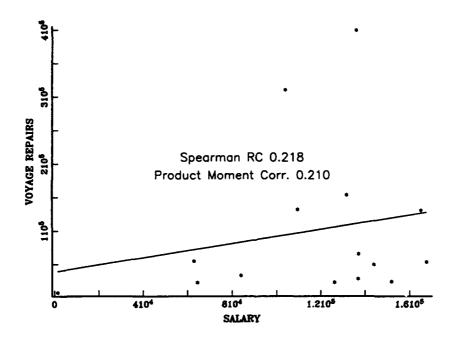


**Figure 59**
**T-ATF 166 Salary vs Voyage Repair Cost Scatter Plot**

ATF TRAINING VS SUBSISTENCE (LESS HI LEVERAGE PT)
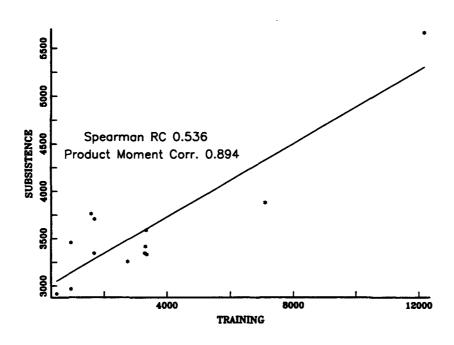


**Figure 60**
**T-ATF 166 Training vs Subsistence Cost Scatter Plot**

87
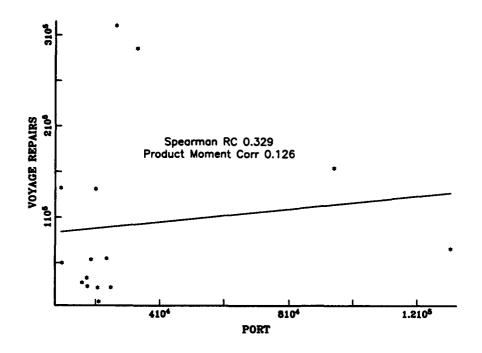
**Figure 61**
**T-ATF 166 Port and Misc. vs Voyage Repair Cost Scatter Plot**

88

# V. THE MODEL

The cost data was obtained from the Accounting Department at MSCPAC. This data was downloaded from the Financial Management Information System (FMIS). The format chosen for download was the Nucleus Report, a monthly report in which the costs incurred for each ship during the month are listed by cost category, with all sub-category line items listed.

## A. COST ANALYSIS MODEL

The spreadsheet model is based on a normal costing system. The actual costs recorded monthly for the seven cost categories will be considered as direct costs since they are easily attributable to each ship using FMIS. Overhead costs are considered indirect costs, and budgeted costs will be used. Figure 62 is an overview of the costing system used for the Cost Simulation Tool.

Each run of the simulation will yield direct cost estimates for each of the seven cost categories used in the model. The simulation software generates a new value for each run as a result of a Monte Carlo simulation which will be discussed in detail in Chapter VI. The simulation generates a value based on the probability distribution assumed for each category. These assumptions are based on the data analysis.

89
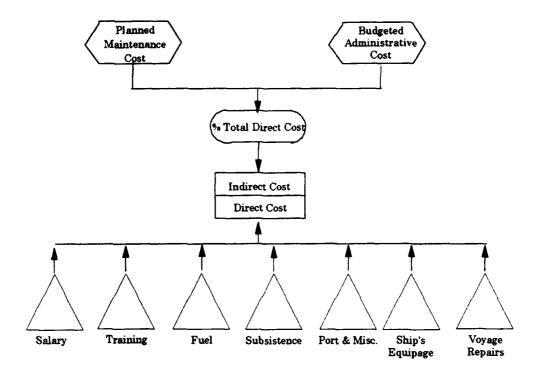
**Figure 62**
**Cost Simulation Model Overview**

The seven direct cost categories are shown as triangles in Figure 62. These costs are represented as random variables for which a new value is produced with each replication of the simulation. The values are then summed by the spreadsheet.

The two indirect cost categories are taken into account in the model as budgeted indirect overhead costs. The value is allocated as a pre-determined percentage of the sum of the direct costs for the replication. The indirect cost value is then summed by the spreadsheet.

90

## B. MICROSOFT EXCEL SPREADSHEET MODEL

The model was designed to capitalize on the Nucleus Report format. Figure 63 is a printout of the T-AO 187 Class Cost Simulation Tool model. The model consists of a Microsoft Excel worksheet where the cost categories are listed and summed for a period of time. The user can choose the period of time desired and the number of ships to sum. The mean monthly amount for each cost category and a figure for the estimated amount for each cost category are listed for the assumed period of time and number of ships.

The mean monthly amount for each category is based on the results of the data analysis covered in detail in Chapter IV. The user can change the mean amount using the edit assumptions button for the category. This allows the user to perform "what if" analysis by increasing or decreasing the mean monthly amount used for the simulation run. When the Edit button is clicked, the Crystal Ball dialog box for the assumption appears. The user must be careful to enter his desired entry in the correct edit box. It is essential that the user review the Crystal Ball user's manual prior to using the spreadsheet model. The underlying distributions should not be changed by the user since these distributions were the result of extensive data analysis.

91

## TAO Time Analysis Sheet

**Salary Assumptions**

| | | |
|---|---|---|
| Mean monthly salary | ▓▓▓▓▓▓▓▓ | Edit |
| Estimated salary | $22,907.55 | |

**Training Assumptions**

| | | |
|---|---|---|
| Mean monthly training expense | ▓▓▓▓▓▓▓▓ | Edit |
| Estimated training cost | $740.80 | |

**Fuel Assumptions**

| | | |
|---|---|---|
| Mean monthly fuel cost | ▓▓▓▓▓▓▓▓ | Edit |
| Estimated fuel cost | $6,808.07 | |

**Subsistence Assumptions**

| | | |
|---|---|---|
| Mean monthly subsistence cost | ▓▓▓▓▓▓▓▓ | Edit |
| Estimated subsistence cost | $641.85 | |

**Port and Misc. Assumptions**

| | | |
|---|---|---|
| Mean monthly port and misc. cost | ▓▓▓▓▓▓▓▓ | Edit |
| Estimated port and miscellaneous cost | $5,026.83 | |

**Ship's Equipage Assumption**

| | | |
|---|---|---|
| Mean monthly ships equipage cost | ▓▓▓▓▓▓▓▓ | Edit |
| Estimated ship's equipage cost | $681.76 | |

**Voyage Repair Assumptions**

| | | |
|---|---|---|
| Mean monthly voyage repair costs | ▓▓▓▓▓▓▓▓ | Edit |
| Estimated voyage repair cost | $3,370.68 | |

**Indirect Overhead Budgeted**

| | | |
|---|---|---|
| Budgeted monthly overhead costs | ▓▓▓▓▓▓▓▓ | Edit |
| Estimated overhead costs | $4,945.05 | |

**Time Assumption**

| | | |
|---|---|---|
| Number of months to forecast for | ▓▓▓▓▓▓▓▓ | Edit |

**Number of Ships Assumption**

| | | |
|---|---|---|
| Number of ships to operate for forecast perio | ▓▓▓▓▓▓▓▓ | Edit |

## Total Cost Forecast          $40,340.57

( Report )          ( Profit )          ( Run )

**Figure 63**
**TAO-187 Class Time Analysis Worksheet**

The user may also edit the assumed value for budgeted overhead costs. This value is not affected by Crystal Ball since it is deterministic. The user may change this value by increasing or decreasing the budgeted overhead by a percentage.

The user may choose a time period to run the simulation for by clicking on the edit button for that assumption in the spreadsheet. The user must then enter a value for time in either days, weeks, or months. If the user does not select a time, the default value for time is one month.

The user may choose the number of ships for the simulation run by clicking on the corresponding edit button. If the user does not provide an input, the default value for number of ships will be one.

If it is desired to input a target operating income (profit), the user must click on the profit button at the bottom of the sheet. This will cause the total cost forecast to include the target income in its total. The target operating income can be in terms of a percentage of total cost or a specified amount depending on the desires of the user. The default setting for profit is zero since this is primarily a cost model.

Once all of the options have been selected by the user, the simulation is started by clicking the run button at the bottom of the sheet. This will cause the simulation to commence using whatever run preferences have been

input into Crystal Ball. If the user desires to change the run preferences, this is done using the pull down menu. The Crystal Ball user's guide contains detailed instructions for changing the dialog box that appears for run preferences.

During the simulation run, a graphical representation of the empirical distribution created by the simulation values for total cost appears on the screen. Figure 64 is a screen capture showing this graph. The graph allows the user to follow the progress of the simulation.

At the completion of the simulation, similar graphs will be displayed for each of the cost categories. This allows the user to analyze each cost category separately. If the user desires a detailed report of the simulation results, he must click the "Report" button at the bottom of the spreadsheet.
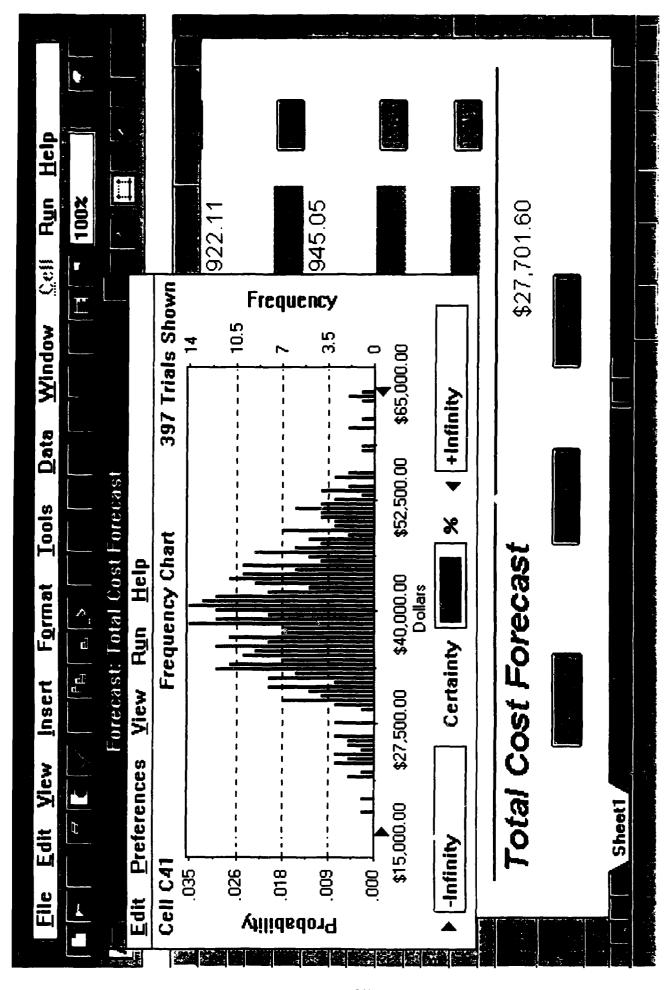
Figure 64
Microsoft Excel Worksheet Screen Print

# VI. SIMULATION METHOD

The purpose of this chapter is to present a brief description of the simulation method used by Crystal Ball, the simulation software used for this model. Crystal Ball uses a probabilistic Monte Carlo method to generate cost values which are summed on a Microsoft Excel spreadsheet.

Problems handled by Monte Carlo methods are of two types, probabilistic or deterministic, depending on whether or not they are directly concerned with the behavior and outcome of random processes. In our case, the chief assumption has been that the direct costs behave according to probability distributions. Each cost category was fitted with a probability distribution in the data analysis performed in Chapter IV.

In the probabilistic Monte Carlo case, the simplest approach is to observe random numbers, chosen in such a way that they directly simulate the random processes of the cost categories, and to infer the desired solution from the behavior of these random numbers. A probability density function is considered to have a total area, enclosed by the curve and the horizontal axis, of unity. A random number with a value between zero and one is generated by the computer. As seen in Figure 65, a point on the curve can be selected where the area beneath the curve accumulated to that point is equal to the value of the random number.

Area = $x$
Random Number = $x$

A = value generated by $x$
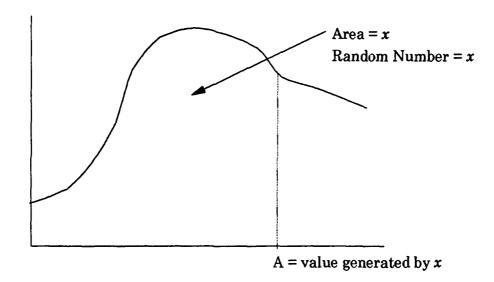
**Figure 65**
**Random number relationship to distribution.**

There is a point, $A$ on the $x$ axis which is the projection of the point on the curve determined by the random number generation. The value of $A$ corresponds to the value to be summed on the spreadsheet.

This process is repeated for each of the cost categories using the probability distribution that was determined by the previous data analysis. After all values are obtained, the spreadsheet is summed to determine the total cost value for the period of time and number of ships concerned.

The advantage of using a computer to perform these tasks is that the process can quickly be replicated to generate a new set of values. With each replication, another value is created for each cost category and the total cost. These values generated for each run are plotted so that a graph is generated which represents an empirical distribution of all the values generated for the

97

simulation. This empirical distribution is also analyzed to determine the mean and other salient characteristics of the values generated by the simulation. The graph can also be examined to find ranges of values which fit our desired certainty for the result.

## A. RANDOM NUMBER GENERATION

An essential feature of Monte Carlo simulation is that at some point a substitution must be made for a random variable using a set of randomly generated values having the statistical characteristics of the random variable. The values that are substituted are called *random numbers* on the basis that they could well have been produced by chance by a suitable random process. As it turns out, the random numbers are not produced in this way, however, this should not affect our use of them. The question is not "Where did these numbers come from?" but rather "Are these numbers correctly distributed?" This question is answered by statistical tests on the random numbers themselves that are beyond the scope of this effort.

When the term random number is used in Monte Carlo simulation, the standardized uniform distribution, $U(0,1)$, is being referred to. In our case, the numbers are being generated by a *pseudo-random* number generation algorithm using a personal computer. The great advantage of this method is that the sequence of pseudorandom numbers can be exactly reproduced for purposes of computational checking.

Crystal Ball makes use of the Lehmer congruential method[1]. This method capitalizes on the architecture of the computer being used. In the case of the personal computers being used, all having 32 bit central processing units (CPU), the Lehmer method can generate a very long sequence of numbers without repeating a number. A 32 bit CPU permits 4.25 billion numbers to be generated without a repetition. This application will never test these limits!

## B. TRANSFORMATION OF PSEUDO-RANDOM NUMBERS

An important question is "How will the pseudo-random numbers from the uniform distribution be used to represent the various probability distributions in this model?" The method used to apply these pseudo-random variables to probability distributions is based on the inverse of the cumulative distribution function:

Suppose that there are uniform $U(0, 1)$ random variables $u_1, u_s, ..., u_n$. If $X_i \sim F(x)$, we can generate $X_i$ by :

$$X_i = F^{-1}(u_i)$$

For $F^{-1}$ to be unique, $F$ must be strictly monotonic.

The significance of this procedure is that random variables from known distributions can be represented by a sequence of uniform pseudo-random numbers if those numbers are "pushed" through the inverse

---

[1]Decisioneering Inc., *Crystal Ball 3.0 Users Manual*, p 38, Decisioneering Inc., Denver, CO.

cumulative distribution function of the known distribution being represented.

## C. CRYSTAL BALL

Crystal Ball has sixteen distributions available for use. For each of these, Crystal Ball pushes a uniform pseudo-random number through the inverse transformation of the selected probability distribution. The result number is a random number from the selected distribution which is used for the present replication.

Figure 66 is a simplified diagram that shows the sequence of events in the Crystal Ball simulation process. After the distributional assumptions are made and the values and parameters are entered for the spreadsheet cells in which the simulation values are generated (referred to by Crystal Ball as the assumption cell), the simulation is started. Next, the random numbers are generated for each assumption cell and the transformations are made.

Once the simulated values for each assumption cell are obtained, the spreadsheet is calculated. The result of this calculation yields the forecast values being sought. Once the forecast values are obtained, the process is repeated until the number of runs entered by the user is reached. On each run, the values for each category of cost are averaged with all of the previous values obtained in the previous runs. This average, or grand mean, is the

**Simulation Cycle**
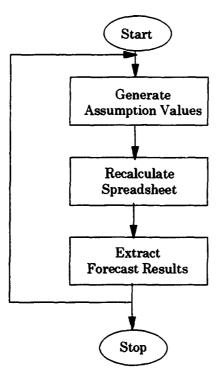


**Figure 66**

**Crystal Ball simulation overview.**

mean value achieved so far for the simulation. At the end of the simulation, the mean value represents the final grand mean value for the entire simulation.

# VII. VALIDATION OF DATA ANALYSIS

The distributional assumptions made in Chapter IV were based on using a partition of the entire data set consisting of two chronological halves. The earlier half of the entire monthly cost data set was used during the data analysis. The later half of the entire data set was next used to validate the assumptions made in Chapter IV. If the distributional assumptions made earlier remain valid, the validation data set should be well fit by the same distributions as those chosen earlier.

An additional check was made by running the Cost Simulation Tool model against several cases of the actual data for randomly selected months. If the data analysis and model are valid, the simulation run values should reasonably represent the actual monthly values.

## A. SUMMARY OF DATA ANALYSIS VALIDATION RESULTS

The validation data sets were analyzed in two ways. First, probability distribution fit plots were made for each cost category to assess the fit of the distribution family found for the corresponding cost category in Chapter IV to the validation data. For these fits, the parameters were estimated using the validation data. A second probability distribution fit was also attempted using the same distribution family and the parameters estimated in Chapter IV.

For every cost category, a good fit to the validation data was obtained with the distribution family used for the corresponding cost category in Chapter IV, but with new parameter estimates. When attempting to fit the validation data with the parameters (using same distributions) found in Chapter IV, the fits obtained were almost always poor. This result caused another look to be taken at the data. Since the sample size was relatively small with only half of the data set, both halves were used to get more precise estimates of the parameters.

This time, the entire data set (filtered as in Chapter IV) for each category was used. Again, two fits were made for each cost category. The first set of plots was made by attempting to fit the data set with both the distribution and parameters from Chapter IV. The second set of plots was made by attempting to fit the data set with only the distribution from Chapter IV, allowing AGSS to automatically use parameters that provide the best fit . In this round of plotting, the fits to the full data set were better than those found with the validation data set when using both the distribution and parameters from Chapter IV as would be expected. Excellent results were achieved when fitting the distribution only to the full data sets. The model was updated to reflect the new results obtained from allowing AGSS to automatically select parameters that best fit the entire

data set. Validation test runs of the model were again made after the changes were made to the model.

Figure 67 shows the original probability distribution plot for T-ATF 166 Class Voyage Repairs which was based on the first half of the original partitioned data set. Figure 68 is the new probability distribution plot for the same cost category using the full data set. It is evident from comparing the two views that the data is equally well fit by the same distribution though the parameters have changed slightly.
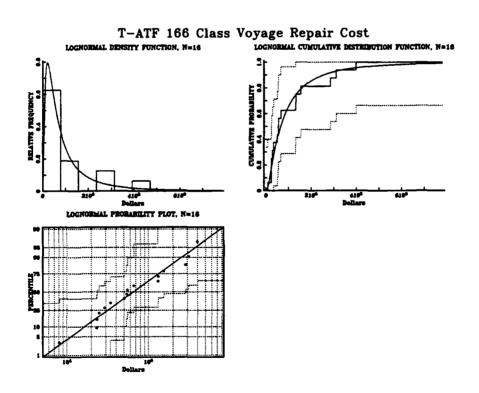


**Figure 67**
**T-ATF 166 Class Voyage Repair Probability Distribution Fit Plot.**
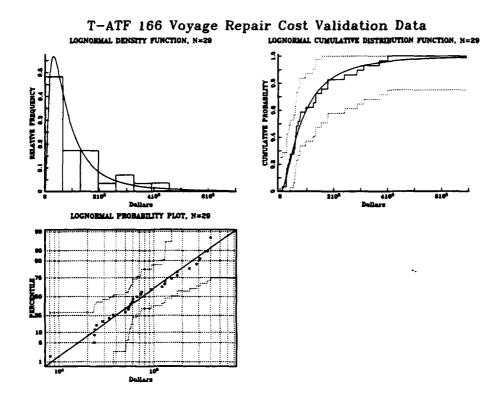
**Figure 68**
**T-ATF 166 Revised Voyage Repair Probability Distribution Fit**

All cost categories were treated in the same way and the results were compared to insure that the two sets of parameters found (Chapter IV and Chapter VII) for each cost category were within a 95 percent confidence interval of each other (based on fitting 95 percent confidence interval to Chapter IV estimates and testing to ensure that Chapter VII estimates fell inside.) All data sets passed this test. A brief summary of the results of the improved distribution fits obtained using the entire data sets is given in Tables 5 and 6.

## Table 5
## T-AO 187 CLASS DATA ANALYSIS VALIDATION RESULTS

| Cost Category | Distribution | Parameters | Mean | Std Dev | Fit Quality |
|---|---|---|---|---|---|
| Salary | Logistic | α=732,000, β=93,164 | 725,644 | 170,648 | Good |
| Training | Weibull | C=1.5962, a=19,615 | 17,689 | 12,008 | Good |
| Fuel | Logistic | α=208,464, β=48,002 | 204,943 | 105,861 | Good |
| Subsistence | Logistic | α=19,230, β=3,579 | 19,342 | 7,758 | Fair |
| Port and Misc. | Weibull | C=1.579, a=220,080 | 196,010 | 141,300 | Fair |
| Ship's Equipage | Weibull | C=0.96836, a=19,117 | 19,492 | 21,107 | Good |
| Voyage Repairs | Gamma | α=0.97569, β=119,370 | 118,400 | 97,767 | Good |

## Table 6
## T-ATF 166 CLASS DATA ANALYSIS VALIDATION RESULTS

| Cost Category | Distribution | Parameters | Mean | Std Dev | Fit Quality |
|---|---|---|---|---|---|
| Salary | Normal | μ=143,830, σ=74,280 | 143,830 | 74,280 | Fair |
| Training | Gamma | α=1.791, β=1,444 | 2,588 | 2,356 | Good |
| Fuel | Weibull | C=1.2947, a=60,386 | 56,961 | 49,056 | Fair |
| Subsistence | Logistic | α=3,424, β=916.2 | 3,642 | 1,902 | Marginal |
| Port and Misc. | Lognormal | μ=10.371, σ=1.0614 | 59,353 | 83,532 | Good |
| Ship's Equipage | Lognormal | μ=7.5506, σ=0.4695 | 4,111 | 5,274 | Good |
| Voyage Repairs | Lognormal | μ=11.236, σ=0.95045 | 115,380 | 108,540 | Good |

## B. VALIDATION RESULTS USING COST SIMULATION MODEL

Once the model parameters were updated as a result of the previous graphical data analysis, the Cost Simulation Tool was run for several cases to compare the results of the simulation to the actual values documented with historical cost data. These validation test runs will be discussed on an individual case basis.

106

## 1. Case 1: T-AO 187 Class Per Diem Rate Comparison

MSCPAC charges their customers for services provided by their ships on a per diem basis. Each ship class has a per diem rate that has been computed by MSC headquarters and is charged to the sponsor for each ship day of use (this per diem rate includes overhead costs and the corresponding model used for the validation runs also included overhead costs). In this case, the T-AO 187 Class model was run for a time period of one ship day. The results were compared against the per diem rate of $60,765 per ship day (FY 94 rate). If the model and data analysis accurately represent the actual costs incurred, the result obtained by running the model for one ship day should be close to the per-diem rate presently being charged.

Figure 69 is a combined empirical probability distribution/frequency plot generated by the simulation. The actual rate of $60,765 falls at the 98th percentile of the empirical distribution. This implies that approximately 98 percent of the time, the actual per diem rate charged will be higher than the result of a simulation run. This result indicates that either the model is predicting lower values than actually necessary to cover operating costs, or the present per diem rate is set too high. Since the per diem rate falls within the values generated by the simulation, the simulation method does seem to generate values similar in

**Forecast: Total Cost Forecast**

**Cell C41**　　　　**Frequency Chart**　　　　**1,989 Trials Shown**

Certainty Range is from -Infinity to $60,765.00 Dollars
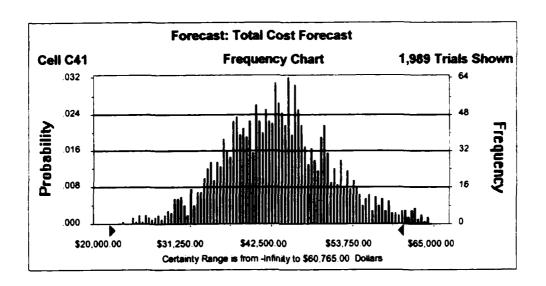
**Figure 69**
**T-AO 187 Class Per Diem Comparison Result.**

value to the per diem rate. The next case, which simulates T-AO 187 direct
costs only, provides some insight.

### 2. Case 2: T-AO 187 Class Historical Data Comparison Method

In this case, the Cost Simulation Tool was run for a period of
one month for one ship. The results of this simulation run were then
compared to the actual monthly costs reported for all ships (data set was
filtered according to the model assumptions). Table 7 shows a randomly
selected actual monthly data set for five ships and indicates how many
standard deviations away from the simulation mean for the cost category
(from Table 5) the selected data was. A better view of the model accuracy is
seen in Figure 70 that compares the actual total costs with the simulation
mean total cost. An interval of one standard deviation of the total cost

108

## Table 7
## T-AO 187 CLASS COMPARISON (ACTUAL VS SIMULATION)

| Cost Category | Ship 1 | #SD | Ship 2 | #SD | Ship 3 | #SD | Ship 4 | #SD | Ship 5 | #SD |
|---|---|---|---|---|---|---|---|---|---|---|
| Salary | 319,177 | -2.382 | 510,241 | -1.292 | 793,971 | 0.383 | 961,026 | 1.370 | 964,076 | 1.388 |
| Training | 11,081 | -0.550 | 17,937 | 0.056 | 3,251 | -1.230 | 51,928 | 2.976 | 22,291 | 0.437 |
| Fuel | 243,552 | 0.365 | 415,734 | 2.610 | 169,215 | -0.564 | 126,511 | -1.114 | 140,132 | -0.938 |
| Subsistence | 29,651 | ˉ47 | 18,699 | -0.110 | 29,651 | 1.647 | 20,461 | 0.173 | 11,948 | -1.193 |
| Port and Misc. | 586,889 | 2.766 | 115,181 | -1.016 | 196,304 | 0.046 | 280,271 | 1.145 | 186,436 | -0.084 |
| Ship's Equipage | 43,601 | 1.082 | 754 | -0.896 | 16,213 | -0.183 | 5,559 | -0.675 | 13,086 | -0.327 |
| Voyage Repairs | 264,090 | 1.490 | 312,193 | 1.565 | 202,048 | 0.667 | 425,969 | 2.492 | 113,970 | -0.051 |
| Total Cost | 1,498,041 | 0.813 | 1,390,739 | 0.349 | 1,410,653 | 0.435 | 1,871,725 | 2.429 | 1,451,939 | 0.614 |
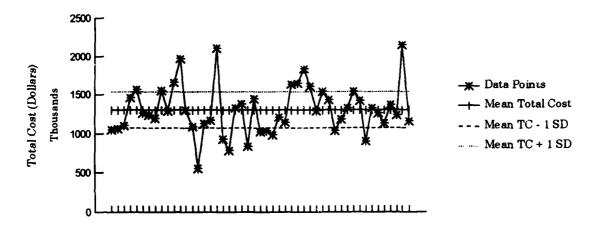


## Figure 70
## T-AO 187 Class Monthly Actual Data Comparison Result.

obtained by the model run above and below the simulation mean is also shown in the plot.

For the monthly data comparison run, the budgeted overhead feature of the model was turned off. This was done because there is no actual overhead data included with the actual cost data as reported by FMIS. The overhead feature is used in the model for estimating the per-diem rate only; when comparing to actual cost categories listed in FMIS, there is no category

109

for budgeted overhead. The actual costs are mostly (69.4%) within a range of one standard deviation of the simulation mean value.

The observation that the model seemed to predict values lower than the actual per-diem rate (as seen in case 1) would imply that the model would predict lower monthly totals than the actual monthly data. This is not seen in case 2. The model values seem to be about the same as the actual values. The actual direct costs are all close to the simulation mean value (with the budgeted overhead feature turned off). This suggests that either the budgeted overhead rate used in the model is lower than what would be needed to make the actual per-diem rate accurate or that the per-diem rate is too high . It appears that the model is predicting direct costs accurately.

### 3. Case 3: T-ATF 166 Class Per Diem Rate Comparison

In this case, the per diem rate being charged for FY 94 is $15,955. Figure 71 is an empirical probability distribution/frequency plot generated by the simulation. This plot indicates that the actual per diem rate falls at the 77.6 percentile. This result indicates that the model predictions are more accurate than those of case one and that the actual per diem rate being charged for this class vessel is closer to the actual costs of operation than was found in case one.
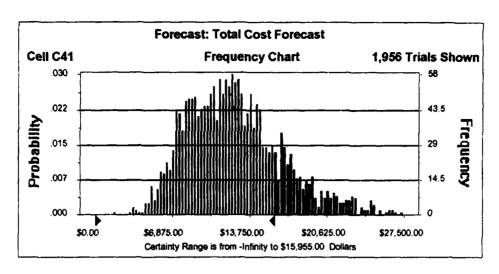
**Figure 71**
**T-ATF 166 Class Per Diem Comparison Result.**

## 4. Case 4: T-ATF 166 Class Historical Data Comparison Method

This is similar to case two discussed in section B above. Table 8 is the actual monthly cost comparison data by cost category for five randomly selected T-ATF 166 class ship-months.

Other than a few isolated instances, the actual data was close to the simulation mean values in every cost category (from Table 6). In most

**Table 8**
**T-ATF 166 CLASS COMPARISON (ACTUAL VS SIMULATION)**

| Cost Category | Ship 1 | #SD | Ship 2 | #SD | Ship 3 | #SD | Ship 4 | #SD | Ship 5 | #SD |
|---|---|---|---|---|---|---|---|---|---|---|
| Salary | 109,516 | -0.535 | 130,546 | -0.024 | 131,804 | -0.224 | 141,973 | -0.081 | 259,192 | 1.559 |
| Training | 12,178 | 5.157 | 1,783 | -0.407 | 481 | -1.104 | 1,756 | -0.422 | 3,519 | 0.522 |
| Fuel | 43,866 | -0.283 | 44,584 | -0.267 | 11,671 | -1.004 | 70,042 | 0.303 | 27,509 | -0.650 |
| Subsistence | 3,459 | -0.078 | 5,771 | 1.406 | 2,919 | -0.424 | 3,441 | -0.089 | 1,361 | -1.424 |
| Port and Misc. | 9,289 | -0.579 | 7,674 | -0.598 | 94,103 | 0.409 | 14,431 | -0.519 | 32,801 | -0.305 |
| Ship's Equipage | 4,159 | 0.030 | 25,619 | 4.527 | 10,486 | 1.354 | 860 | -0.662 | 9,191 | 1.084 |
| Voyage Repairs | 132,368 | 0.156 | 122,597 | 0.067 | 154,183 | 0.355 | 23,782 | -0.835 | 285,624 | 1.554 |
| Total Cost | 314,835 | -0.464 | 338,574 | -0.304 | 405,647 | 0.149 | 256,285 | -0.860 | 619,197 | 1.590 |

cases, the values were within one standard deviation of the simulation mean value. Figure 72 is a plot that compares the actual total costs with the simulation mean total cost. The values are mostly (72.4%) within one standard deviation of the simulation mean. The simulation produces a reasonable representation of the data from this ship class .

## 5. Summary of Model Comparison Results

The results seen in cases three and four suggest that the model concept is correct but that the overhead assumptions in the T-AO 187 class model need to be examined for accuracy. In both cases, the majority of the actual results were within one standard deviation of the simulation mean. As the primary focus has been the simulation of actual values, the model is viable in that respect.
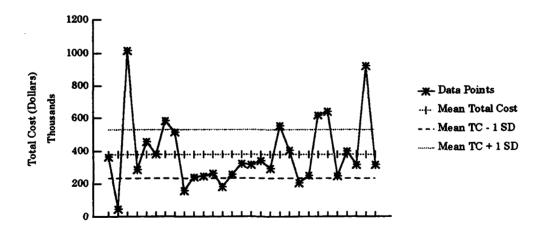


**Figure 72**
**T-ATF 166 Class Monthly Actual Data Comparison Result.**

# VIII. CONCLUSIONS

The original intent of this project was to develop a tool that would enable MSCPAC to perform reasonably accurate "what if" cost analyses for the ships they own and operate. The Cost Simulation Tool was developed to fill that requirement. As in any computer application, the expected results from the product are only as good as the inputs to the program. This is the famous "*garbage in = garbage out*" principle of computing.

A qualified success has been achieved with the Cost Simulation Tool. The success is in the ability of the simulation to accurately forecast direct ship operating costs. Unfortunately, in the area of indirect overhead costs, more work is needed, which is beyond the scope of this thesis. The overhead estimate used in the Cost Simulation Tool model was a best guess figure provided by MSCPAC operations department and is not documented anywhere. To correct this problem would require a cost analysis to determine the extent of the overhead costs for MSC. The total overhead costs might then be allocated as a percentage of total direct costs.

The original problem statement expressed by the MSCPAC comptroller included some degree of certainty about the costs in the areas of infrastructure, overhead, and chartering services. As this project has unfolded, the concerns expressed by the comptroller in the area of direct costs for the ships owned and operated by MSC were real, but not as serious as

113

perceived. This analysis has shown that the cost information available from the FMIS was accurate but cumbersome to access and use. The Cost Simulation Tool has been demonstrated to quickly and reliably produce operating cost estimates for the direct costs. The overhead costs; however, are not precisely known, and when the rough estimate is included in the cost simulation model, the accuracy of the model suffers.

The problem of imprecise overhead estimates is not unique to MSC. Since the advent of DBOF, most Department of Defense activities subjected to this mode of funding have suffered from an inability to accurately forecast overhead amounts to be included in billing. The subject of overhead cost analysis is wide open and needs further examination in the case of MSC.

The direct costs of operating the ships are reported in great detail and can be directly traced to their source in an economically feasible manner. The Financial Management Information System (FMIS) maintains records of every voucher written by the ships. Clearly, the record keeping is complete in this area. The Cost Simulation Tool exploits this historical database to produce accurate direct cost results by simulation.

A project of enormous use to MSC management would be to enhance FMIS to include a query system that could allow the managers to access data and manipulate it from their Microsoft Excel spreadsheets. This would permit the data of interest to be displayed in tables and graphs without the

114

necessity of mastering the complexities of the FMIS system itself. The combination of the Cost Simulation Tool and the query system would give the managers great flexibility in manipulating the data and analyzing the outcome of various scenarios. The two systems could work together to give management a better understanding of the operating costs of their business.

Another analysis could be conducted to determine the overhead costs and their allocation at MSC. The result of this future analysis could be incorporated into the Cost Simulation Tool to improve its accuracy. The Cost Simulation Tool could then be used to calculate per diem rates that would reflect actual MSC costs with much greater accuracy.

Without following up to determine the overhead costs and their drivers, use of the Cost Simulation Tool is limited to estimating direct costs only. The Cost Simulation Tool was shown to produce accurate results in the area of direct costs, but the total cost accuracy is subject to the rough estimate of budgeted overhead. This is because the overhead cost calculation is, at present, based on a "best guess" approximation of overhead costs.

The question is, "What has been accomplished by the Cost Simulation Tool?" First, the data analysis has shown that the operating cost data is indeed random in nature and can be represented by probability distributions. Second, once the probability distributions have been determined, the costs can be simulated using the Monte Carlo simulation method.

These results, when combined, give the cost accountant a new tool to use in forecasting the costs of doing business. No longer tied to regression alone; through simulation, the cost accountant can see the entire range of behavior for the costs of concern. This ability can give forecasts much greater accuracy than in the past.

# APPENDIX A

## Figure 2 Data Table

ANALYSIS OF LOGISTIC DISTRIBUTION FIT

```
DATA        : OSALXY14
SELECTION   : ALL
X AXIS LABEL: Dollars x 964,075
SAMPLE SIZE : 25
CENSORING   : NONE
FREQUENCIES : 1
EST. METHOD : MAXIMUM L KELIHOOD
CONF METHOD : ASYMPTOTIC NORMAL APPROXIMATION
```

```
                     CONF. INTERVALS    COVARIANCE MATRIX OF
                       (95 PERCENT)      PARAMETER ESTIMATES
PARAMETER  ESTIMATE  LOWER    UPPER       ALPHA      BETA
ALPHA      0.71841  0.65081 0.78602    0.0011892  0
BETA       0.099551 0.06691 0.13219    0          0.00027722
LOG LIKELIHOOD FUNCTION AT MLE = 8.0678
```

```
              SAMPLE     FITTED              GOODNESS OF FIT TESTS
MEAN     :   0.70757    0.71841       CHI-SQUARE :  2.2654
STD DEV :    0.17471    0.18057          DEG FREED:  1
SKEWNESS:   -0.34818    0                  SIGNIF  :  0.1323
KURTOSIS:    2.2898     4.2           KOLM-SMIRN :  0.11668
• BASED ON MIDPOINTS OF FINITE INTERVALS     SIGNIF :  0.88546
                                       CRAMER-V M :  0.063255
PERCENTILES  SAMPLE     FITTED           SIGNIF   :   > .15
       5:    0.40366   0.42529       ANDER-DARL :  0.4681
      10:    0.4413    0.49968          SIGNIF  :   > .15        •-
      25:    0.61702   0.60804
      50:    0.75042   0.71841
      75:    0.82356   0.82778       KS, AD, AND CV SIGNIF. LEVELS NOT
      90:    0.89774   0.93715       EXACT WITH ESTIMATED PARAMETERS.
      95:    0.99684   1.0115
                                     NOTE:  A SMALL SIGNIFICANCE LEVEL
                                     (EG. P≤.01) INDICATES LACK OF FIT
```

```
           CHI-SQUARE GOODNESS OF FIT TABLE

LOWER    UPPER    OBS   EXP     O-E      ((O-E)•2)÷E
-INF.    0.61714   7   6.6388  0.36123   0.019655
0.61714  0.71999   3   5.9605 -2.9605    1.4705
0.71999  0.82285   8   5.9157  2.0843    0.73436
0.82285  +INF.     7   6.485   0.515     0.040899
TOTAL             25  25                 2.2654
```


## Figure 7 Data Table

ANALYSIS OF WEIBULL DISTRIBUTION FIT

```
DATA        : OTRAIN07
SELECTION   : ALL
X AXIS LABEL: Dollars
SAMPLE SIZE : 25
CENSORING   : NONE
FREQUENCIES : 1
EST. METHOD : MAXIMUM LIKELIHOOD
CONF METHOD : ASYMPTOTIC NORMAL APPROXIMATION
```

```
                     CONF. INTERVALS    COVARIANCE MATRIX OF
                       (95 PERCENT)      PARAMETER ESTIMATES
PARAMETER  ESTIMATE  LOWER     UPPER       C         a
C (SHAPE)    1.5615  1.0715   2.0515    0.062478  2.5573E2
a (SCALE)  24864    18307    31422      255.73    1.1189E7
LOG LIKELIHOOD FUNCTION AT MLE = -272.27
```

```
              SAMPLE     FITTED              GOODNESS OF FIT TESTS
MEAN     :  22411      22346          CHI-SQUARE :  1.2157
STD DEV :   14578      14822             DEG FREED:  2
SKEWNESS:    0.81912    1.0009            SIGNIF  :  0.54454
KURTOSIS:    2.364      1.4288        KOLM-SMIRN :  0.09619
• BASED ON MIDPOINTS OF FINITE INTERVALS     SIGNIF :  0.97484
                                       CRAMER-V M :  0.031256
PERCENTILES  SAMPLE     FITTED           SIGNIF   :   > .15
       5:    3251.4    3710.8        ANDER-DARL :  0.24434
      10:    5637      5884.1           SIGNIF  :   > .15
      25:   11395     11196
      50:   17937     19662         KS, AD, AND CV SIGNIF. LEVELS NOT
      75:   27089     30650         EXACT WITH ESTIMATED PARAMETERS.
      90:   45251     42418
      95:   49853     50204         NOTE:  A SMALL SIGNIFICANCE LEVEL
                                    (EG. P≤.01) INDICATES LACK OF FIT
```

```
           CHI-SQUARE GOODNESS OF FIT TABLE

LOWER    UPPER    OBS   EXP     O-E       ((O-E)•2)÷E
-INF.    8473.4    3   4.2473 -1.2473     0.36631
8473.4  16847      6   6.3229 -0.32287    0.016487
16847   25420      7   5.5502  1.4498     0.37869
25420   33894      3   3.9425 -0.94252    0.22533
33894   +INF.      6   4.937   1.063      0.22886
TOTAL             25  25                  1.2157
```

117

# APPENDIX A

## Figure 11 Data Table

ANALYSIS OF LOGISTIC DISTRIBUTION FIT

```
DATA          :  OFUEL14
SELECTION     :  ALL
X AXIS LABEL:    Dollars X 534,166
SAMPLE SIZE   :  25
CENSORING     :  NONE
FREQUENCIES   :  1
EST. METHOD   :  MAXIMUM LIKELIHOOD
CONF METHOD   :  ASYMPTOTIC NORMAL APPROXIMATION
```

|            |          | CONF. INTERVALS (95 PERCENT) | | COVARIANCE MATRIX OF PARAMETER ESTIMATES | |
|------------|----------|--------|--------|-----------|----------|
| PARAMETER  | ESTIMATE | LOWER  | UPPER  | ALPHA     | BETA     |
| ALPHA      | 0.37821  | 0.30979 | 0.44664 | 0.0012183 | 0       |
| BETA       | 0.10076  | 0.067723 | 0.1338 | 0         | 0.000284 |

LOG LIKELIHOOD FUNCTION AT MLE = 6.0708

|          | SAMPLE  | FITTED  |
|----------|---------|---------|
| MEAN   : | 0.38495 | 0.37821 |
| STD DEV : | 0.20786 | 0.18276 |
| SKEWNESS: | 0.72848 | 0 |
| KURTOSIS: | 5.1271 | 4.2 |

• BASED ON MIDPOINTS OF FINITE INTERVALS

| GOODNESS OF FIT TESTS | |
|---------------|---------|
| KOLM-SMIRN : | 0.15724 |
| SIGNIF  : | 0.56679 |
| CRAMER-V M : | 0.10111 |
| SIGNIF  : | > .15 |
| ANDER-DARL : | 0.71454 |
| SIGNIF  : | > .15 |

| PERCENTILES | SAMPLE | FITTED |
|-------------|--------|--------|
| 5: | 0.0063472 | 0.081531 |
| 10: | 0.072774 | 0.15682 |
| 25: | 0.31678 | 0.26752 |
| 50: | 0.38108 | 0.37821 |
| 75: | 0.46914 | 0.48891 |
| 90: | 0.52469 | 0.59961 |
| 95: | 0.77829 | 0.6749 |

KS, AD, AND CV SIGNIF. LEVELS NOT
EXACT WITH ESTIMATED PARAMETERS.

NOTE:  A SMALL SIGNIFICANCE LEVEL
(EG. P≤.01) INDICATES LACK OF FIT

## Figure 15 Data Table

ANALYSIS OF LOGISTIC DISTRIBUTION FIT

```
DATA          :  OSUBS09
SELECTION     :  ALL
X AXIS LABEL:    Dollars X 40177
SAMPLE SIZE   :  25
CENSORING     :  NONE
FREQUENCIES   :  1
EST. METHOD   :  MAXIMUM LIKELIHOOD
CONF METHOD   :  ASYMPTOTIC NORMAL APPROXIMATION
```

|            |          | CONF. INTERVALS (95 PERCENT) | | COVARIANCE MATRIX OF PARAMETER ESTIMATES | |
|------------|----------|--------|--------|-----------|----------|
| PARAMETER  | ESTIMATE | LOWER  | UPPER  | ALPHA     | BETA     |
| ALPHA      | 19486    | 16957  | 22016  | 1.6651E6  | 0.0000E0 |
| BETA       | 3725.1   | 2503.7 | 4946.4 | 0.0000E0  | 3.8815E5 |

LOG LIKELIHOOD FUNCTION AT MLE = -257.48

|          | SAMPLE  | FITTED  |
|----------|---------|---------|
| MEAN   : | 19824 | 19486 |
| STD DEV : | 7997.5 | 6756.5 |
| SKEWNESS: | 0.44477 | 0 |
| KURTOSIS: | 5.0097 | 4.2 |

• BASED ON MIDPOINTS OF FINITE INTERVALS

| GOODNESS OF FIT TESTS | |
|---------------|---------|
| KOLM-SMIRN : | 0.21565 |
| SIGNIF  : | 0.19532 |
| CRAMER-V M : | 0.27207 |
| SIGNIF  : | > .15 |
| ANDER-DARL : | 1.5124 |
| SIGNIF  : | > .15 |

| PERCENTILES | SAMPLE | FITTED |
|-------------|--------|--------|
| 5: | 5870.5 | 8518 |
| 10: | 13177 | 11301 |
| 25: | 17996 | 15394 |
| 50: | 19749 | 19486 |
| 75: | 20461 | 23579 |
| 90: | 29651 | 27671 |
| 95: | 38236 | 30454 |

KS, AD, AND CV SIGNIF. LEVELS NOT
EXACT WITH ESTIMATED PARAMETERS.

NOTE:  A SMALL SIGNIFICANCE LEVEL
(EG. P≤.01) INDICATES LACK OF FIT

# APPENDIX A

## Figure 19 Data Table

ANALYSIS OF WEIBULL DISTRIBUTION FIT

```
DATA         :  OPORTO6
SELECTION    :  ALL
X AXIS LABEL :  Dollars
SAMPLE SIZE  :  25
CENSORING    :  NONE
FREQUENCIES  :  1
EST. METHOD  :  MAXIMUM LIKELIHOOD
CONF METHOD  :  ASYMPTOTIC NORMAL APPROXIMATION
```

|  | | CONF. INTERVALS (95 PERCENT) | | COVARIANCE MATRIX OF PARAMETER ESTIMATES | |
|---|---|---|---|---|---|
| PARAMETER | ESTIMATE | LOWER | UPPER | C | a |
| C (SHAPE) | 2.7352E0 | 1.8876E0 | 3.5828E0 | 0.18694 | 1.7119E3 |
| a (SCALE) | 1.7101E5 | 1.4530E5 | 1.9673E5 | 1711.9 | 1.7205E8 |

LOG LIKELIHOOD FUNCTION AT MLE = -310.59

| | SAMPLE | FITTED |
|---|---|---|
| MEAN   : | 1.5257E5 | 1.5215E5 |
| STD DEV : | 8.1177E4 | 8.0071E4 |
| SKEWNESS: | 8.4080E-2 | 8.3849E-1 |
| KURTOSIS: | 2.8129E0 | 1.2909E0 |

* BASED ON MIDPOINTS OF FINITE INTERVALS

| PERCENTILES | SAMPLE | FITTED |
|---|---|---|
| 5: | 6.2816E4 | 5.7733E4 |
| 10: | 7.4355E4 | 7.5114E4 |
| 25: | 1.1518E5 | 1.0845E5 |
| 50: | 1.5866E5 | 1.4957E5 |
| 75: | 1.9297E5 | 1.9271E5 |
| 90: | 2.2008E5 | 2.3199E5 |
| 95: | 2.6501E5 | 2.5541E5 |

```
            GOODNESS OF FIT TESTS
CHI-SQUARE :  0.4684
  DEG FREED:  2
  SIGNIF   :  0.7912
KOLM-SMIRN :  0.088328
  SIGNIF   :  0.98844
CRAMER-V M :  0.026164
  SIGNIF   :  > .15
ANDER-DARL :  0.1888
  SIGNIF   :  > .15

KS, AD, AND CV SIGNIF LEVELS NOT
EXACT WITH ESTIMATED PARAMETERS.

NOTE:  A SMALL SIGNIFICANCE LEVEL
(EG. P≤.01) INDICATES LACK OF FIT
```

CHI-SQUARE GOODNESS OF FIT TABLE

| LOWER | UPPER | OBS | EXP | O-E | ((O-E)*2)÷E |
|---|---|---|---|---|---|
| -INF. | 8.7443E4 | 4 | 3.6893 | 0.31067 | 0.026161 |
| 8.7443E4 | 1.3116E5 | 5 | 5.9031 | -0.9031 | 0.13818 |
| 1.3116E5 | 1.7489E5 | 8 | 6.7734 | 1.2267 | 0.22214 |
| 1.7489E5 | 2.1861E5 | 5 | 5.1033 | -0.10328 | 0.0020894 |
| 2.1861E5 | +INF. | 3 | 3.531 | -0.53096 | 0.079841 |
| TOTAL | | 25 | 25 | | 0.4684 |

## Figure 23 Data Table

ANALYSIS OF WEIBULL DISTRIBUTION FIT

```
DATA         :  OEQUIP05
SELECTION    :  ALL
X AXIS LABEL :  Dollars
SAMPLE SIZE  :  25
CENSORING    :  NONE
FREQUENCIES  :  1
EST. METHOD  :  MAXIMUM LIKELIHOOD
CONF METHOD  :  ASYMPTOTIC NORMAL APPROXIMATION
```

|  | | CONF. INTERVALS (95 PERCENT) | | COVARIANCE MATRIX OF PARAMETER ESTIMATES | |
|---|---|---|---|---|---|
| PARAMETER | ESTIMATE | LOWER | UPPER | C | a |
| C (SHAPE) | 1.0816 | 0.73862 | 1.4247 | 0.030616 | 2.1536E2 |
| a (SCALE) | 21037 | 13039 | 29035 | 215.36 | 1.6646E7 |

LOG LIKELIHOOD FUNCTION AT MLE = -273.04

| | SAMPLE | FITTED |
|---|---|---|
| MEAN   : | 20456 | 20414 |
| STD DEV : | 17708 | 18890 |
| SKEWNESS: | 1.0125 | 1.2115 |
| KURTOSIS: | 3.1307 | 1.6612 |

* BASED ON MIDPOINTS OF FINITE INTERVALS

| PERCENTILES | SAMPLE | FITTED |
|---|---|---|
| 5: | 753.5 | 1350.2 |
| 10: | 1958 | 2826.8 |
| 25: | 9331.6 | 6648.7 |
| 50: | 14067 | 14991 |
| 75: | 28376 | 28453 |
| 90: | 46797 | 45484 |
| 95: | 50740 | 58012 |

```
            GOODNESS OF FIT TESTS
CHI-SQUARE :  1.4944
  DEG FREED:  1
  SIGNIF   :  0.22154
KOLM-SMIRN :  0.10514
  SIGNIF   :  0.94511
CRAMER-V M :  0.044199
  SIGNIF   :  > .15
ANDER-DARL :  0.29121
  SIGNIF   :  > .15

KS, AD, AND CV SIGNIF. LEVELS NOT
EXACT WITH ESTIMATED PARAMETERS.

NOTE:  A SMALL SIGNIFICANCE LEVEL
(EG. P≤.01) INDICATES LACK OF FIT
```

CHI-SQUARE GOODNESS OF FIT TABLE

| LOWER | UPPER | OBS | EXP | O-E | ((O-E)*2)÷E |
|---|---|---|---|---|---|
| -INF. | 11001 | 8 | 9.7754 | -1.7754 | 0.32244 |
| 11001 | 22001 | 9 | 6.4733 | 2.5267 | 0.98622 |
| 22001 | 33002 | 3 | 3.8408 | -0.84081 | 0.18407 |
| 33002 | +INF. | 5 | 4.9105 | 0.089504 | 0.0016314 |
| TOTAL | | 25 | 25 | | 1.4944 |

119

# APPENDIX A

## Figure 27 Data Table

ANALYSIS OF GAMMA DISTRIBUTION FIT

```
DATA         :  OVREPOS
SELECTION    :  ALL
X AXIS LABEL :  Dollars
SAMPLE SIZE  :  25
CENSORING    :  NONE
FREQUENCIES  :  1
EST. METHOD  :  MAXIMUM LIKELIHOOD
CONF METHOD  :  ASYMPTOTIC NORMAL APPROXIMATION
```

|           |          | CONF. INTERVALS (95 PERCENT) | |          | COVARIANCE MATRIX OF PARAMETER ESTIMATES | |
|-----------|----------|---------|---------|---|---------|---------|
| PARAMETER | ESTIMATE | LOWER   | UPPER   |   | ALPHA   | BETA    |
| ALPHA     | 8.3755E-1 | 0.43484 | 1.2402E0 | | 0.042177 | -6.4306E3 |
| BETA      | 1.2770E5 | 45473   | 2.0993E5 | | -6430.6 | 1.7593E9 |

LOG LIKELIHOOD FUNCTION AT MLE = -339.34

|          | SAMPLE   | FITTED   |
|----------|----------|----------|
| MEAN     | 1.0885E5 | 1.0885E5 |
| STD DEV  | 9.7542E4 | 1.1887E5 |
| SKEWNESS | 9.3881E-1 | 2.1854E0 |
| KURTOSIS | 3.4970E0 | 1.0184E1 |

* BASED ON MIDPOINTS OF FINITE INTERVALS

| PERCENTILES | SAMPLE | FITTED |
|-------------|--------|--------|
| 5:  | 3.3061E3 | 3.3838E3 |
| 10: | 6.1825E3 | 7.8749E3 |
| 25: | 2.1116E4 | 2.5245E4 |
| 50: | 8.3043E4 | 6.8580E4 |
| 75: | 1.8079E5 | 1.4805E5 |
| 90: | 2.2253E5 | 2.5720E5 |
| 95: | 2.2788E5 | 3.4129E5 |

```
          GOODNESS OF FIT TESTS
CHI-SQUARE :  0.76465
  DEG FREED:  1
  SIGNIF   :  0.38188
KOLM-SMIRN :  0.11317
  SIGNIF   :  0.90803
CRAMER-V M :  0.072187
  SIGNIF   :  > .15
ANDER-DARL :  0.47451
  SIGNIF   :  > .15

KS, AD, AND CV SIGNIF. LEVELS NOT
EXACT WITH ESTIMATED PARAMETERS.

NOTE:  A SMALL SIGNIFICANCE LEVEL
(EG. P≤.0°) INDICATES LACK OF FIT
```

CHI-SQUARE GOODNESS OF FIT TABLE

| LOWER    | UPPER    | OBS | EXP    | O-E     | ((O-E)*2)+E |
|----------|----------|-----|--------|---------|-------------|
| -INF.    | 6.3591E4 | 11  | 11.924 | -0.9241 | 0.071616 |
| 6.3591E4 | 1.2718E5 | 5   | 5.6027 | -0.6027 | 0.064833 |
| 1.2718E5 | 1.9077E5 | 3   | 3.1206 | -0.12063 | 0.0046631 |
| 1.9077E5 | +INF.    | 6   | 4.3526 | 1.6474  | 0.62354 |
| TOTAL    |          | 25  | 25     |         | 0.76465 |

## Figure 31 Data Table

ANALYSIS OF NORMAL DISTRIBUTION FIT

```
DATA         :  TSALRY02
SELECTION    :  ALL
X AXIS LABEL :  Dollars
SAMPLE SIZE  :  16
CENSORING    :  NONE
FREQUENCIES  :  1
EST. METHOD  :  MAXIMUM LIKELIHOOD
CONF METHOD  :  EXACT
```

|           |          | CONF. INTERVALS (95 PERCENT) | |          | COVARIANCE MATRIX OF PARAMETER ESTIMATES | |
|-----------|----------|---------|---------|---|---------|---------|
| PARAMETER | ESTIMATE | LOWER   | UPPER   |   | MU      | SIGMA   |
| MU        | 1.2388E5 | 93687   | 1.5408E5 | | 1.8813E8 | 0.0000E0 |
| SIGMA     | 5.4864E4 | 41857   | 8.7699E4 | | 0.0000E0 | 9.4066E7 |

LOG LIKELIHOOD FUNCTION AT MLE = -197.31

|          | SAMPLE   | FITTED   |
|----------|----------|----------|
| MEAN     | 1.2388E5 | 1.2388E5 |
| STD DEV  | 5.6664E4 | 5.4864E4 |
| SKEWNESS | 1.1985E-1 | 0.0000E0 |
| KURTOSIS | 4.1197E0 | 3.0000E0 |

* BASED ON MIDPOINTS OF FINITE INTERVALS

| PERCENTILES | SAMPLE | FITTED |
|-------------|--------|--------|
| 5:  | 1.3497E3 | 3.3620E4 |
| 10: | 6.2848E4 | 5.3563E4 |
| 25: | 9.3969E4 | 8.6895E4 |
| 50: | 1.3393E5 | 1.2388E5 |
| 75: | 1.4800E5 | 1.6087E5 |
| 90: | 1.6765E5 | 1.9421E5 |
| 95: | 2.5919E5 | 2.1415E5 |

```
          GOODNESS OF FIT TESTS
KOLM-SMIRN :  0.15001
  SIGNIF   :  0.86422
CRAMER-V M :  0.075901
  SIGNIF   :  > .15
ANDER-DARL :  0.47923
  SIGNIF   :  > .15

KS, AD, AND CV SIGNIF. LEVELS NOT
EXACT WITH ESTIMATED PARAMETERS.

NOTE:  A SMALL SIGNIFICANCE LEVEL
(EG. P≤.01) INDICATES LACK OF FIT
```

120

# APPENDIX A

## Figure 35 Data Table

ANALYSIS OF GAMMA DISTRIBUTION FIT

```
DATA         :  TTRAIN07
SELECTION    :  ALL
X AXIS LABEL :  Dollars
SAMPLE SIZE  :  16
CENSORING    :  NONE
FREQUENCIES  :  1
EST. METHOD  :  MAXIMUM LIKELIHOOD
CONF METHOD  :  ASYMPTOTIC NORMAL APPROXIMATION
```

|           |          | CONF. INTERVALS (95 PERCENT) | | COVARIANCE MATRIX OF PARAMETER ESTIMATES | |
|-----------|----------|-------|-------|---------|---------|
| PARAMETER | ESTIMATE | LOWER | UPPER | ALPHA | BETA |
| ALPHA     | 1.3843   | 0.51501 | 2.2537 | 0.19665 | -3.1855E2 |
| BETA      | 2242.5   | 552.64  | 3932.4 | -318.55 | 7.4307E5 |

LOG LIKELIHOOD FUNCTION A: MLE = -160.29

|          | SAMPLE | FITTED | GOODNESS OF FIT TESTS | |
|----------|--------|--------|-----------|---------|
| MEAN     | 3104.4 | 3104.4 | KOLM-SMIRN : | 0.20696 |
| STD DEV  | 2944.1 | 2638.5 | SIGNIF    : | 0.4996 |
| SKEWNESS | 1.9927 | 1.6998 | CRAMER-V M : | 0.08997 |
| KURTOSIS | 6.8261 | 7.3342 | SIGNIF    : | > .15 |

* BASED ON MIDPOINTS OF FINITE INTERVALS

```
ANDER-DARL :  0.4847
SIGNIF     :  > .15
```

| PERCENTILES | SAMPLE | FITTED |
|-------------|--------|--------|
| 5:          | 181.26 | 317.11 |
| 10:         | 481.26 | 545.54 |
| 25:         | 1259.4 | 1185.1 |
| 50:         | 3018   | 2397.2 |
| 75:         | 3344   | 4267.4 |
| 90:         | 7105.3 | 6597.3 |
| 95:         | 12178  | 8308.3 |

KS, AD, AND CV SIGNIF. LEVELS NOT
EXACT WITH ESTIMATED PARAMETERS.

NOTE:  A SMALL SIGNIFICANCE LEVEL
(EG. P≤.01) INDICATES LACK OF FIT

## Figure 39 Data Table

ANALYSIS OF WEIBULL DISTRIBUTION FIT

```
DATA         :  TFUEL08
SELECTION    :  ALL
X AXIS LABEL :  Dollars
SAMPLE SIZE  :  16
CENSORING    :  NONE
FREQUENCIES  :  1
EST. METHOD  :  MAXIMUM LIKELIHOOD
CONF METHOD  :  ASYMPTOTIC NORMAL APPROXIMATION
```

|           |          | CONF. INTERVALS (95 PERCENT) | | COVARIANCE MATRIX OF PARAMETER ESTIMATES | |
|-----------|----------|-------|-------|---------|---------|
| PARAMETER | ESTIMATE | LOWER | UPPER | C | α |
| C (SHAPE) | 1.7563   | 1.0625 | 2.45  | 0.12524 | 8.1824E2 |
| α (SCALE) | 49874    | 35236  | 64511 | 818.24  | 5.5747E7 |

LOG LIKELIHOOD FUNCTION AT MLE = -184.17

|          | SAMPLE | FITTED | GOODNESS OF FIT TESTS | |
|----------|--------|--------|-----------|---------|
| MEAN     | 44424  | 44409  | KOLM-SMIRN : | 0.13903 |
| STD DEV  | 26742  | 26105  | SIGNIF    : | 0.91656 |
| SKEWNESS | 0.43147 | 0.9342 | CRAMER-V M : | 0.032114 |
| KURTOSIS | 2.3588 | 1.3818 | SIGNIF    : | > .15 |

* BASED ON MIDPOINTS OF FINITE INTERVALS

```
ANDER-DARL :  0.211
SIGNIF     :  > .15
```

| PERCENTILES | SAMPLE | FITTED |
|-------------|--------|--------|
| 5:          | 6565.9 | 9191.6 |
| 10:         | 11672  | 13848  |
| 25:         | 23229  | 24535  |
| 50:         | 43152  | 40480  |
| 75:         | 62744  | 60068  |
| 90:         | 89412  | 80189  |
| 95:         | 96866  | 93151  |

KS, AD, AND CV SIGNIF. LEVELS NOT
EXACT WITH ESTIMATED PARAMETERS.

NOTE:  A SMALL SIGNIFICANCE LEVEL
(EG. P≤.01) INDICATES LACK OF FIT

# APPENDIX A

## Figure 43 Data Table

ANALYSIS OF LOGISTIC DISTRIBUTION FIT

```
DATA          :  TSUBS07
SELECTION     :  ALL
X AXIS LABEL:    Dollars
SAMPLE SIZE   :  16
CENSORING     :  NONE
FREQUENCIES   :  1
EST. METHOD   :  MAXIMUM LIKELIHOOD
CONF METHOD   :  ASYMPTOTIC NORMAL APPROXIMATION
```

```
                    CONF. INTERVALS    COVARIANCE MATRIX OF
                     (95 PERCENT)       PARAMETER ESTIMATES
PARAMETER  ESTIMATE  LOWER    UPPER        ALPHA     BETA
ALPHA       3432.7   2757.3  4108.1      1.187E5     0
BETA        795.66   469.56  1121.8      0.000E0   27670
LOG LIKELIHOOD FUNCTION AT MLE = -140.47
```

```
            SAMPLE     FITTED                    GOODNESS OF FIT TESTS
MEAN    :   3669.8     3432.7              KOLM-SMIRN :  0.23775
STD DEV :   1894       1443.2                SIGNIF   :  0.32626
SKEWNESS:   1.9387     0                   CRAMER-V M :  0.2226
KURTOSIS:   7.6241     4.2                   SIGNIF   :   > .15
* BASED ON MIDPOINTS OF FINITE INTERVALS   ANDER-DARL :  1.2804
                                             SIGNIF   :   > .15

PERCENTILES  SAMPLE    FITTED
      5:     1039.3    1090               KS, AD, AND CV SIGNIF. LEVELS NOT
     10:     1360.7    1684.5             EXACT WITH ESTIMATED PARAMETERS.
     25:     3113.7    2558.6
     50:     3383.1    3432.7             NOTE:  A SMALL SIGNIFICANCE LEVEL
     75:     3735.7    4306.8             (EG. P≤.01) INDICATES LACK OF FIT
     90:     5662.2    5181
     95:     9659.3    5775.5
```

## Figure 47 Data Table

ANALYSIS OF LOGNORMAL DISTRIBUTION FIT

```
DATA          :  TPORT08
SELECTION     :  ALL
X AXIS LABEL:    Dollars
SAMPLE SIZE   :  16
CENSORING     :  NONE
FREQUENCIES   :  1
EST. METHOD   :  MAXIMUM LIKELIHOOD
CONF METHOD   :  EXACT
```

```
                    CONF. INTERVALS    COVARIANCE MATRIX OF
                     (95 PERCENT)       PARAMETER ESTIMATES
PARAMETER  ESTIMATE  LOWER    UPPER        MU       SIGMA
MU          10.253   9.727   10.778     0.057004    0
SIGMA       0.95502  0.72861  1.5266    0          0.028502
LOG LIKELIHOOD FUNCTION AT MLE = -186.01
```

```
            SAMPLE     FITTED                    GOODNESS OF FIT TESTS
MEAN    :   55316      44741              KOLM-SMIRN :  0.28052
STD DEV :   98294      54603                SIGNIF   :  0.16115
SKEWNESS:   3.036      5.4791            CRAMER-V M :  0.26189
KURTOSIS:   11.176     84.858              SIGNIF   :   > .15
* BASED ON MIDPOINTS OF FINITE INTERVALS   ANDER-DARL :  1.3477
                                             SIGNIF   :   > .15

PERCENTILES  SAMPLE    FITTED
      5:     9.2892E3  5.8922E3           KS, AD, AND CV SIGNIF. LEVELS NOT
     10:     9.6902E3  8.3375E3           EXACT WITH ESTIMATED PARAMETERS.
     25:     1.7410E4  1.4894E4
     50:     2.0855E4  2.8356E4           NOTE:  A SMALL SIGNIFICANCE LEVEL
     75:     2.9579E4  5.3985E4           (EG. P≤.01) INDICATES LACK OF FIT
     90:     1.3051E5  9.6441E4
     95:     4.0299E5  1.3646E5
```

# APPENDIX A

## Figure 51 Data Table

ANALYSIS OF LOGNORMAL DISTRIBUTION FIT

```
DATA          :  TEQUIP05
SELECTION     :  ALL
X AXIS LABEL:   Dollars
SAMPLE SIZE   :  16
CENSORING     :  NONE
FREQUENCIES   :  1
EST. METHOD   :  MAXIMUM LIKELIHOOD
CONF METHOD   :  EXACT
```

|  | | CONF. INTERVALS (95 PERCENT) | | COVARIANCE MATRIX OF PARAMETER ESTIMATES | |
|---|---|---|---|---|---|
| PARAMETER | ESTIMATE | LOWER | UPPER | MU | SIGMA |
| MU | 7.4952 | 6.7489 | 8.2416 | 0.11493 | 0 |
| SIGMA | 1.3561 | 1.0346 | 2.1676 | 0 | 0.057466 |

LOG LIKELIHOOD FUNCTION AT MLE = −147.5

|  | SAMPLE | FITTED | | GOODNESS OF FIT TESTS | |
|---|---|---|---|---|---|
| MEAN | 3682.3 | 4512.8 | KOLM-SMIRN : | 0.13769 | |
| STD DEV : | 3941.3 | 10379 | SIGNIF : | 0.92204 | |
| SKEWNESS: | 1.0952 | 19.066 | CRAMER-V M : | 0.050035 | |
| KURTOSIS: | 2.9511 | 2178.4 | SIGNIF : | > .15 | |

- BASED ON MIDPOINTS OF FINITE INTERVALS   ANDER-DARL : 0.31448
                                            SIGNIF : > .15

| PERCENTILES | SAMPLE | FITTED |
|---|---|---|
| 5: | 138.08 | 193.29 |
| 10: | 185.64 | 316.43 |
| 25: | 666.93 | 721.24 |
| 50: | 2541 | 1799.4 |
| 75: | 5279.6 | 4489.3 |
| 90: | 10486 | 10233 |
| 95: | 12568 | 16751 |

KS, AD, AND CV SIGNIF. LEVELS NOT EXACT WITH ESTIMATED PARAMETERS.

NOTE: A SMALL SIGNIFICANCE LEVEL (EG. P≤.01) INDICATES LACK OF FIT

## Figure 55 Data Table

ANALYSIS OF LOGNORMAL DISTRIBUTION FIT

```
DATA          :  TVREP06
SELECTION     :  ALL
X AXIS LABEL:   Dollars
SAMPLE SIZE   :  16
CENSORING     :  NONE
FREQUENCIES   :  1
EST. METHOD   :  MAXIMUM LIKELIHOOD
CONF METHOD   :  EXACT
```

|  | | CONF. INTERVALS (95 PERCENT) | | COVARIANCE MATRIX OF PARAMETER ESTIMATES | |
|---|---|---|---|---|---|
| PARAMETER | ESTIMATE | LOWER | UPPER | MU | SIGMA |
| MU | 11.082 | 10.495 | 11.67 | 0.071287 | 0 |
| SIGMA | 1.068 | 0.81479 | 1.7071 | 0 | 0.035644 |

LOG LIKELIHOOD FUNCTION AT MLE = −201.07

|  | SAMPLE | FITTED | | GOODNESS OF FIT TESTS | |
|---|---|---|---|---|---|
| MEAN : | 1.1140E5 | 1.1501E5 | KOLM-SMIRN : | 0.12223 | |
| STD DEV : | 1.1962E5 | 1.6779E5 | SIGNIF : | 0.9706 | |
| SKEWNESS: | 1.3067E0 | 7.4826E0 | CRAMER-V M : | 0.052749 | |
| KURTOSIS: | 3.4176E0 | 1.8342E2 | SIGNIF : | > .15 | |

- BASED ON MIDPOINTS OF FINITE INTERVALS   ANDER-DARL : 0.3337
                                            SIGNIF : > .15

| PERCENTILES | SAMPLE | FITTED |
|---|---|---|
| 5: | 8.0831E3 | 1.1219E4 |
| 10: | 2.3280E4 | 1.6541E4 |
| 25: | 2.6833E4 | 3.1647E4 |
| 50: | 5.4715E4 | 6.5019E4 |
| 75: | 1.4328E5 | 1.3358E5 |
| 90: | 3.1072E5 | 2.5558E5 |
| 95: | 3.9986E5 | 3.7681E5 |

KS, AD, AND CV SIGNIF. LEVELS NOT EXACT WITH ESTIMATED PARAMETERS.

NOTE: A SMALL SIGNIFICANCE LEVEL (EG. P≤.01) INDICATES LACK OF FIT

# APPENDIX A

## Figure 67 Data Table

ANALYSIS OF LOGNORMAL DISTRIBUTION FIT

```
DATA         : TVREP06
SELECTION    : ALL
X AXIS LABEL: Dollars
SAMPLE SIZE  : 16
CENSORING    : NONE
FREQUENCIES  : 1
EST. METHOD  : MAXIMUM LIKELIHOOD
CONF METHOD  : EXACT
```

|           |          | CONF. INTERVALS (95 PERCENT) | | COVARIANCE MATRIX OF PARAMETER ESTIMATES | |
|-----------|----------|--------|--------|----------|----------|
| PARAMETER | ESTIMATE | LOWER  | UPPER  | MU       | SIGMA    |
| MU        | 11.082   | 10.495 | 11.67  | 0.071287 | 0        |
| SIGMA     | 1.068    | 0.81479| 1.7071 | 0        | 0.035644 |

LOG LIKELIHOOD FUNCTION AT MLE = -201.07

|          | SAMPLE   | FITTED   |
|----------|----------|----------|
| MEAN   : | 1.1140E5 | 1.1501E5 |
| STD DEV: | 1.1962E5 | 1.6779E5 |
| SKEWNESS:| 1.3067E0 | 7.4826E0 |
| KURTOSIS:| 3.4176E0 | 1.8342E2 |

• BASED ON MIDPOINTS OF FINITE INTERVALS

| PERCENTILES | SAMPLE   | FITTED   |
|-------------|----------|----------|
| 5:          | 8.0831E3 | 1.1219E4 |
| 10:         | 2.3280E4 | 1.6541E4 |
| 25:         | 2.6833E4 | 3.1647E4 |
| 50:         | 5.4715E4 | 6.5019E4 |
| 75:         | 1.4328E5 | 1.3358E5 |
| 90:         | 3.1072E5 | 2.5558E5 |
| 95:         | 3.9986E5 | 3.7681E5 |

GOODNESS OF FIT TESTS

```
KOLM-SMIRN :  0.12223
    SIGNIF :  0.9706
CRAMER-V M :  0.052749
    SIGNIF :  > .15
ANDER-DARL :  0.3337
    SIGNIF :  > .15
```

KS, AD, AND CV SIGNIF. LEVELS NOT EXACT WITH ESTIMATED PARAMETERS.

NOTE:  A SMALL SIGNIFICANCE LEVEL (EG. P≤.01) INDICATES LACK OF FIT

## Figure 68 Data Table

ANALYSIS OF LOGNORMAL DISTRIBUTION FIT

```
DATA         : FULLVAL2[:7]
SELECTION    : ALL
X AXIS LABEL: Dollars
SAMPLE SIZE  : 29
CENSORING    : NONE
FREQUENCIES  : 1
EST. METHOD  : MAXIMUM LIKELIHOOD
CONF METHOD  : EXACT
```

|           |          | CONF. INTERVALS (95 PERCENT) | | COVARIANCE MATRIX OF PARAMETER ESTIMATES | |
|-----------|----------|--------|--------|---------|----------|
| PARAMETER | ESTIMATE | LOWER  | UPPER  | MU      | SIGMA    |
| MU        | 11.236   | 10.868 | 11.604 | 0.03115 | 0        |
| SIGMA     | 0.95045  | 0.7676 | 1.3083 | 0       | 0.015575 |

LOG LIKELIHOOD FUNCTION AT MLE = -365.51

|          | SAMPLE   | FITTED   |
|----------|----------|----------|
| MEAN   : | 1.1538E5 | 1.1904E5 |
| STD DEV: | 1.0854E5 | 1.4423E5 |
| SKEWNESS:| 1.3233E0 | 5.4130E0 |
| KURTOSIS:| 3.6926E0 | 8.2422E1 |

• BASED ON MIDPOINTS OF FINITE INTERVALS

| PERCENTILES | SAMPLE   | FITTED   |
|-------------|----------|----------|
| 5:          | 2.3280E4 | 1.5865E4 |
| 10:         | 2.3548E4 | 2.2412E4 |
| 25:         | 3.7556E4 | 3.9926E4 |
| 50:         | 6.6150E4 | 7.5778E4 |
| 75:         | 1.5418E5 | 1.4382E5 |
| 90:         | 3.1072E5 | 2.5622E5 |
| 95:         | 3.6854E5 | 3.6194E5 |

GOODNESS OF FIT TESTS

```
CHI-SQUARE :  1.8769
  DEG FREED:  1
    SIGNIF :  0.17069
KOLM-SMIRN :  0.095754
    SIGNIF :  0.95304
CRAMER-V M :  0.042478
    SIGNIF :  > .15
ANDER-DARL :  0.28923
    SIGNIF :  > .15
```

KS, AD, AND CV SIGNIF. LEVELS NOT EXACT WITH ESTIMATED PARAMETERS.

NOTE:  A SMALL SIGNIFICANCE LEVEL (EG. P≤.01) INDICATES LACK OF FIT

CHI-SQUARE GOODNESS OF FIT TABLE

| LOWER    | UPPER    | OBS | EXP    | O-E    | ((O-E)•2)÷E |
|----------|----------|-----|--------|--------|-------------|
| -INF.    | 6.5297E4 | 14  | 12.695 | 1.3047 | 0.13408     |
| 6.5297E4 | 1.3059E5 | 5   | 8.085  | -3.085 | 1.1771      |
| 1.3059E5 | 1.9589E5 | 5   | 3.6135 | 1.3865 | 0.53198     |
| 1.9589E5 | +INF.    | 5   | 4.6061 | 0.39386| 0.033677    |
| TOTAL    |          | 29  | 29     |        | 1 8769      |

# APPENDIX B

```
' Edit Time Assumption Dialog Box Module

' LCDR Terry Redman, USN
' Naval Postgraduate School
' 1994

'***********************************************************************
'
' Button19_Click Macro
'
' Worksheet Edit Button Actions
'
Sub Button19_Click()

    DialogSheets("Dialog1").Show
    DialogSheets("Dialog1").OptionButtons("Option Button 20") = xlOn
    DialogSheets("Dialog1").EditBoxes("Edit Box 23").Text = "1"
    DialogSheets("Dialog1").EditBoxes("Edit Box 14").Text = " "
    DialogSheets("Dialog1").EditBoxes("Edit Box 19").Text = " "

End Sub

'***********************************************************************
'
' Button2_Click Macro
'
' OK Button actions
'
Sub Button2_Click()

    If DialogSheets("Dialog1").OptionButtons("Option Button 20").Value = xlOn Then W
orksheets("Sheet1").Cells(36, 3).Value = DialogSheets("Dialog1").EditBoxes("Edit Box
 23").Text
    If DialogSheets("Dialog1").OptionButtons("Option Button 11").Value = xlOn Then W
orksheets("Sheet1").Cells(36, 3).Value = ((DialogSheets("Dialog1").EditBoxes("Edit B
ox 14").Text) * (0.032967033))
    If DialogSheets("Dialog1").OptionButtons("Option Button 15").Value = xlOn Then W
orksheets("Sheet1").Cells(36, 3).Value = ((DialogSheets("Dialog1").EditBoxes("Edit B
ox 19").Text) * (0.2307692308))

End Sub

'***********************************************************************
'
' Button3_Click Macro
'
' Cancel button msg
'
Sub Button3_Click()

    MsgBox "Are you sure you want to Cancel?", vbYesNo

End Sub

'***********************************************************************
'
' EditBox14_Change Macro
'
' Days Edit Box Actions
'
```

```
Sub EditBox14_Change()

    DialogSheets("Dialog1").OptionButtons("Option Button 11") = xlOn
    DialogSheets("Dialog1").OptionButtons("Option Button 20") = xlOff
    DialogSheets("Dialog1").OptionButtons("Option Button 15") = xlOff
    MsgBox "Are you sure you want to make this change?", vbYesNo

End Sub

'***********************************************************************
'
' EditBox19_Change Macro
'
' Weeks Edit Box Actions
'
Sub EditBox19_Change()

    DialogSheets("Dialog1").OptionButtons("Option Button 15") = xlOn
    DialogSheets("Dialog1").OptionButtons("Option Button 11") = xlOff
    DialogSheets("Dialog1").OptionButtons("Option Button 20") = xlOff
    MsgBox "Are you sure you want to make this change?", vbYesNo

End Sub

'***********************************************************************
'
' EditBox23_Change Macro
'
' Months Edit Box Actions
'
Sub EditBox23_Change()

    DialogSheets("Dialog1").OptionButtons("Option Button 20") = xlOn
    DialogSheets("Dialog1").OptionButtons("Option Button 11") = xlOff
    DialogSheets("Dialog1").OptionButtons("Option Button 15") = xlOff
    MsgBox "Are you sure you want to make this change?", vbYesNo

End Sub

'***********************************************************************
'
' OptionButton11_Click Macro
'
' Days Option Button Action
'
Sub OptionButton11_Click()

    DialogSheets("Dialog1").OptionButtons("Option Button 20") = .xlOff
    DialogSheets("Dialog1").OptionButtons("Option Button 15") = xlOff

End Sub

'***********************************************************************
'
' OptionButton15_Click Macro
'
' Weeks Option Button Action
'
Sub OptionButton15_Click()
```

```
    DialogSheets("Dialog1").OptionButtons("Option Button 20") = xlOff
    DialogSheets("Dialog1").OptionButtons("Option Button 11") = xlOff

End Sub

'**************************************************************************
'
' OptionButton20_Click Macro
'
' Months Option Button Action
'
Sub OptionButton20_Click()

    DialogSheets("Dialog1").OptionButtons("Option Button 11") = xlOff
    DialogSheets("Dialog1").OptionButtons("Option Button 15") = xlOff

End Sub
```

# APPENDIX B

```
' Run Dialog Box Module

' LCDR Terry Redman, USN
' Naval Postgraduate School
' 1994

'
' Button23_Click Macro
'
' Worksheet Button Actions
'
Sub Button23_Click()

    DialogSheets("Dialog4").Show
    DialogSheets("Dialog4").EditBoxes("Edit Box 6").Text = "2000"

End Sub

'
' Button2_Click Macro
'
' OK Button Actions (Reserved)
'
Sub Button2_Click()

End Sub

'
' Button3_Click Macro
'
' Cancel button msg
'
Sub Button3_Click()

    MsgBox "Are you sure you want to Cancel?", vbYesNo

End Sub
```

# APPENDIX B

```
' Edit Budgeted Overhead Dialog Box Module

' LCDR Terry Redman, USN
' Naval Postgraduate School
' 1994
'********************************************************************************
'
' Button27_Click Macro
'
' Worksheet Button Actions
'
Sub Button27_Click()

    DialogSheets("Dialog2").Show
    DialogSheets("Dialog2").OptionButtons("Option Button 5").Value = xlOn
    DialogSheets("Dialog2").EditBoxes("Edit Box 8").Text = "0"
    DialogSheets("Dialog2").EditBoxes("Edit Box 12").Text = "0"

End Sub

'********************************************************************************
'
' Button2_Click Macro
'
' OK Button Actions (Reserved)
'
Sub Button2_Click()

    If DialogSheets("Dialog2").OptionButtons("Option Button 5").Value = xlOn Then Wo
rksheets("Sheet1").Cells(32, 3).Value = (((DialogSheets("Dialog2").EditBoxes("Edit B
ox 8").Text / 100) + 1) * Worksheets("Sheet1").Cells(32, 3).Value)
    If DialogSheets("Dialog2").OptionButtons("Option Button 9").Value = xlOn Then Wo
rksheets("Sheet1").Cells(32, 3).Value = ((1 - (DialogSheets("Dialog2").EditBoxes("Ed
it Box 12").Text / 100)) * Worksheets("Sheet1").Cells(32, 3).Value)

End Sub

'********************************************************************************
'
' Button3_Click Macro
'
' Cancel button msg
'
Sub Button3_Click()

    MsgBox "Are you sure you want to Cancel?", vbYesNo

End Sub

'********************************************************************************
'
' OptionButton5_Click Macro
'
'
Sub OptionButton5_Click()

    DialogSheets("Dialog2").OptionButtons("Option Button 9").Value = xlOff

End Sub
```

# APPENDIX B

```
'*************************************************************************
'
' OptionButton9_Click Macro
'
'
Sub OptionButton9_Click()

    DialogSheets("Dialog2").OptionButtons("Option Button 5").Value = xlOff

End Sub
```

# APPENDIX B

```
' Edit Number Of Ships Dialog Box Module

' LCDR Terry Redman, USN
' Naval Postgraduate School
' 1994

'*************************************************************************
'
' Button20_Click Macro
'
' Worksheet Button Actions
'
Sub Button20_Click()

    DialogSheets("Dialog3").Show
    DialogSheets("Dialog3").EditBoxes("Edit Box 6").Text = "1"

End Sub

'*************************************************************************
'
' Dialog3_Button2_Click Macro
'
' OK Button Actions
'
Sub Dialog3_Button2_Click()

        Worksheets("Sheet1").Cells(39, 3).Value = DialogSheets("Dialog3").EditBoxes(
"Edit Box 6").Text

End Sub

'*************************************************************************
'
' Button3_Click Macro
'
' Cancel button msg
'
Sub Button3_Click()

    MsgBox "Are you sure you want to Cancel?", vbYesNo

End Sub
```

# APPENDIX B

```
' Crystal Ball Macro Module

' LCDR Terry Redman, USN
' Naval Postgraduate School
' 1994

'*******************************************************************************
    '
    ' AccessAssumption Macro
    ' Macro recorded 4/10/94 by Terry Redman
    '
    '
        Sub AccessAssumption()

            Application.Run Macro:=Range("CB.DefineAssum")

        End Sub

'*******************************************************************************
    ' CreateReport Macro
    ' Macro recorded 4/10/94 by Terry Redman
    '
    '
        Sub CreateReport()

            Application.Run Macro:=Range("CB.CreateRpt")

        End Sub

'*******************************************************************************
    ' RunReplications Macro
    ' Macro recorded 4/10/94 by Terry Redman
    '
    '
        Sub RunReplications()

            Application.Run Macro:=Range("CB.Run")

        End Sub

'*******************************************************************************
    '
    ' ResetRun Macro
    ' Macro recorded 4/10/94 by Terry Redman
    '
    '
        Sub ResetRun()

            Application.Run Macro:=Range("CB.Reset")

        End Sub
```

# References

1. IBM Corporation, <u>A Graphical Statistical System (AGSS):Operators Manual</u>, IBM Research, Yorktown Heights, NY, 1993.

2. Lewis, P.A.W., Orav, E.J., <u>Simulation Methodology for Statisticians, Operations Analysts, and Engineers: Volume I</u>, Wadsworth & Brooks/Cole Advanced Books and Software, Pacific Grove, CA,1989

3. Military Sealift Command; <u>MSCPAC Operations Department QMB</u>, October 1993.

4. Young, Douglas , "Complexities, Impact Of Overhead", U.S. Army Comptroller Office, <u>Resource Management</u>, 1st Quarter, 1994

5. Decisioneering Inc., <u>Crystal Ball Version 3.0</u>, Decisioneering Inc, Denver, CO, 1993.

# Bibliography

Anthony, Robert N.; Hekiman, James S.;Operations Cost Control, Irwin
Corporation, New York, 1969.

Chambers, John M.; Cleveland, Wiliam S.; Kleiner, Beat; Tukey, Paul A.;
Graphical Methods For Data Analysis, Wadsworth International Group,
Belmont, CA, 1983.

Chen, Gordon K.C.; Kaczka, Eugene E.; Operations and Systems Analysis: A
Simulation Approach, Allyn and Bacon, Boston, 1973.

Decisioneering Inc., Crystal Ball Version 3.0, Decisioneering Inc,Denver, CO, 1993.

du Toit, S.H.C.; Steyn, A.G.W.; Stumpf, R.H.; Graphical Exploratory Data Analysis,
Springer-Verlag, New York, 1986.

Freedman, David; Pisani, Robert; Purves, Roger; Adhikari,Ani; Statistics, W.W.
Norton and Company, New York, 1991.

Gibbons, J.D., Chakraborti, S., Nonparametric Statistical Inference, Marcel
Dekker, Inc., New York, 1992

Hammersly, J.M., Handscomb, D.C.; Monte Carlo Methods,Methuen and Co.,
London, UK, 1964.

Horngren, C.T., Foster, G., Datar, S.M., Cost Accounting: A Management
Emphasis, Prentice-Hall, Englewood Cliffs, NJ, 1994

IBM Corporation, A Graphical Statistical System (AGSS): Operators Manual, IBM
Research, Yorktown Heights, NY, 1993.

Kelton, W. David, Law, Averill M.; Simulation Modeling and Analysis,
McGraw-Hill, New York, 1991.

Lewis, P.A.W., Orav, E.J., Simulation Methodology for Statisticians, Operations
Analysts, and Engineers: Volume I, Wadsworth & Brooks/Cole Advanced
Books and Software, Pacific Grove, CA,1989

Microsoft Corporation, Excel 5.0 Users Guide, Microsoft Corporation, Redmond,
WA, 1994.

Military Sealift Command; MSCPAC Operations Department QMB, October 1993.

# Bibliography

Office of the Secretary of Defense, Cost Analysis Improvement Group, <u>Operating and Support Cost Estimating Guide</u>,Government Printing Office, Washington, DC, 1992.

Sutherland, John W., <u>Systems Analysis, Administration, and Architecture</u>, Van Nostrand and Reinhold, New York, 1975.

Wall, Francis J., <u>Statistical Data Analysis Handbook</u>, McGraw-Hill, New York, 1986.

Young, Douglas , "Complexities, Impact Of Overhead", U.S. Army Comptroller Office, <u>Resource Management</u>, 1st Quarter, 1994

# INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center      2
   Cameron Station
   Alexandria, VA 22304-6145

2. Library, Code 0142      2
   Naval Postgraduate School
   Monterey, CA 93943-5000

3. Defense Logistics Studies Information Exchange      2
   U.S. Army Logistics Management Center
   Fort Lee, VA 23801

4. Commander      2
   Military Sealift Command
   Washington Navy Yard, Bldg 210
   901 M Street SE
   Washington, DC 20398-5540

5. Commander      2
   Military Sealift Command, Pacific
   Bldg 310, Naval Supply Center
   Oakland, CA 94625-5010

6. Commander      2
   Military Sealift Command, Atlantic
   Military Ocean Terminal, Bldg 42
   Bayonne, NJ 07002-5399

7. Dr. Dan Boger      1
   Naval Postgraduate School
   Code SM/Bo, Ingersoll Hall
   Monterey, CA 93943-5000

8. Dr. William Kemple      1
   Naval Postgraduate School
   Code OR/Ke, Glasgow Hall
   Monterey, CA 93943-5000

9. Dr. Shu S. Liao      1
   Naval Postgraduate School
   Code SM/Lc, Ingersoll
   Monterey, CA 93943-5000